

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number  
**WO 03/014327 A2**

(51) International Patent Classification<sup>7</sup>: **C12N**

(21) International Application Number: **PCT/US02/26966**

(22) International Filing Date: **9 August 2002 (09.08.2002)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:  
60/310,847 9 August 2001 (09.08.2001) US  
60/336,049 19 November 2001 (19.11.2001) US  
60/338,692 11 December 2001 (11.12.2001) US  
10/171,468 14 June 2002 (14.06.2002) US

(71) Applicant (*for all designated States except US*):  
**MENDEL BIOTECHNOLOGY, INC.** [US/US]; 21375  
Cabot Boulevard, Hayward, CA 94545 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **REUBER, T.,**  
**Lynne** [US/US]; 1115 S. Grant Street, San Mateo, CA  
94402 (US). **RIECHMANN, Jose Luis** [ES/US]; 518  
S. El Molino Avenue, #308, Pasadena, CA 94101 (US).  
**HEARD, Jacqueline, E.** [US/US]; 810 Guilford Av-  
enue, San Mateo, CA 94402 (US). **JIANG, Cai-Zhong**  
[US/US]; 34495 Heathrow Terrace, Fremont, CA 94555  
(US). **ADAM, Luc, J.** [CA/US]; 25800 Industrial Boule-  
vard, Apt. L403, Hayward, CA 94545 (US). **DUBELL,**  
**Arnold, N.** [US/US]; 14857 Wake Avenue, San Leandro,  
CA 94578 (US). **RATCLIFFE, Oliver** [GB/US]; 814 East  
21st Street, Oakland, CA 94606 (US). **PINEDA, Omaira**

[CO/US]; 4060 9th Place, Vero Beach, FL 32960 (US).  
**YU, Guo, Liang** [US/US]; 242 Gravatt Drive, Berkeley,  
CA 94705-1531 (US). **BROUN, Pierre, E.** [FR/US]; 921  
Sunnybrae Boulevard, San Mateo, CA 94402 (US).

(74) Agents: **WARD, Michael, R.** et al.; Morrison & Foerster  
LLP, 425 Market Street, San Francisco, CA 94105-2482  
(US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,  
VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,  
TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *without international search report and to be republished  
upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: **STRESS-RELATED POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS**

(57) Abstract: The invention relates to plant transcription factor polypeptides, polynucleotides that encode them, homologs from a variety of plant species, and methods of using the polynucleotides and polypeptides to produce transgenic plants having advantageous properties compared to a reference plant. Sequence information related to these polynucleotides and polypeptides can also be used in bioinformatic search methods and is also disclosed.

WO 03/014327 A2

## **STRESS-RELATED POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS**

This application claims the benefit of US Provisional Application No. 60/310,847, filed August 9, 2001, US Provisional Application No. 60/336,049, filed December 5, 2001, US Provisional Application No. 60/338,692, filed December 11, 2001, and US Non-provisional Application No. 10/171,468, filed June 14, 2002, the entire contents of which are hereby incorporated by reference.

### **FIELD OF THE INVENTION**

This invention relates to the field of plant biology. More particularly, the present invention pertains to compositions and methods for phenotypically modifying a plant.

### **INTRODUCTION**

A plant's traits, such as its biochemical, developmental, or phenotypic characteristics, may be controlled through a number of cellular processes. One important way to manipulate that control is through transcription factors - proteins that influence the expression of a particular gene or sets of genes. Transformed and transgenic plants that comprise cells having altered levels of at least one selected transcription factor, for example, possess advantageous or desirable traits. Strategies for manipulating traits by altering a plant cell's transcription factor content can therefore result in plants and crops with commercially valuable properties. Applicants have identified polynucleotides encoding transcription factors, developed numerous transgenic plants using these polynucleotides, and have analyzed the plants for a variety of important traits. In so doing, applicants have identified important polynucleotide and polypeptide sequences for producing commercially valuable plants and crops as well as the methods for making them and using them. Other aspects and embodiments of the invention are described below and can be derived from the teachings of this disclosure as a whole.

#### **Background of the Invention**

Transcription factors (TFs) can modulate gene expression, either increasing or decreasing (inducing or repressing) the rate of transcription. This modulation results in differential levels of gene expression at various developmental stages, in different

tissues and cell types, and in response to different exogenous (e.g., environmental) and endogenous stimuli throughout the life cycle of the organism.

Because transcription factors are key controlling elements of biological pathways, altering the expression levels of one or more transcription factors can change entire biological pathways in an organism. For example, manipulation of the levels of selected transcription factors may result in increased expression of economically useful proteins or metabolic chemicals in plants or to improve other agriculturally relevant characteristics. Conversely, blocked or reduced expression of a transcription factor may reduce biosynthesis of unwanted compounds or remove an undesirable trait. Therefore, manipulating transcription factor levels in a plant offers tremendous potential in agricultural biotechnology for modifying a plant's traits.

The present invention provides novel transcription factors useful for modifying a plant's phenotype in desirable ways.

#### **Summary of the Invention**

In a first aspect, the invention relates to a recombinant polynucleotide comprising a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a polypeptide comprising a polypeptide sequence selected from those of the Sequence Listing, SEQ ID NOs: 2 to 2N, where  $N = 2-123$ , or those listed in Table 4, or a complementary nucleotide sequence thereof; (b) a nucleotide sequence encoding a polypeptide comprising a variant of a polypeptide of (a) having one or more, or between 1 and about 5, or between 1 and about 10, or between 1 and about 30, conservative amino acid substitutions; (c) a nucleotide sequence comprising a sequence selected from those of SEQ ID NOs: 1 to  $(2N - 1)$ , where  $N = 2-123$ , or those included in Table 4, or a complementary nucleotide sequence thereof; (d) a nucleotide sequence comprising silent substitutions in a nucleotide sequence of (c); (e) a nucleotide sequence which hybridizes under stringent conditions over substantially the entire length of a nucleotide sequence of one or more of: (a), (b), (c), or (d); (f) a nucleotide sequence comprising at least 10 or 15, or at least about 20, or at least about 30 consecutive nucleotides of a sequence of any of (a)-(e), or at least 10 or 15, or at least about 20, or at least about 30 consecutive nucleotides outside of a region encoding a conserved domain of any of (a)-(e); (g) a

nucleotide sequence comprising a subsequence or fragment of any of (a)-(f), which subsequence or fragment encodes a polypeptide having a biological activity that modifies a plant's characteristic, functions as a transcription factor, or alters the level of transcription of a gene or transgene in a cell; (h) a nucleotide sequence having at least 31% sequence identity to a nucleotide sequence of any of (a)-(g); (i) a nucleotide sequence having at least 60%, or at least 70 %, or at least 80 %, or at least 90 %, or at least 95 % sequence identity to a nucleotide sequence of any of (a)-(g) or a 10 or 15 nucleotide, or at least about 20, or at least about 30 nucleotide region of a sequence of (a)-(g) that is outside of a region encoding a conserved domain; (j) a nucleotide sequence that encodes a polypeptide having at least 31% sequence identity to a polypeptide listed in Table 4, or the Sequence Listing; (k) a nucleotide sequence which encodes a polypeptide having at least 60%, or at least 70 %, or at least 80%, or at least 90 %, or at least 95 % sequence identity to a polypeptide listed in Table 4, or the Sequence Listing; and (l) a nucleotide sequence that encodes a conserved domain of a polypeptide having at least 85%, or at least 90%, or at least 95%, or at least 98% sequence identity to a conserved domain of a polypeptide listed in Table 4, or the Sequence Listing. The recombinant polynucleotide may further comprise a constitutive, inducible, or tissue-specific promoter operably linked to the nucleotide sequence. The invention also relates to compositions comprising at least two of the above-described polynucleotides.

In a second aspect, the invention comprises an isolated or recombinant polypeptide comprising a subsequence of at least about 10, or at least about 15, or at least about 20, or at least about 30 contiguous amino acids encoded by the recombinant or isolated polynucleotide described above, or comprising a subsequence of at least about 8, or at least about 12, or at least about 15, or at least about 20, or at least about 30 contiguous amino acids outside a conserved domain.

In a third aspect, the invention comprises an isolated or recombinant polynucleotide that encodes a polypeptide that is a paralog of the isolated polypeptide described above. In one aspect, the invention is a paralog which, when expressed in *Arabidopsis*, modifies a trait of the *Arabidopsis* plant.



In a fourth aspect, the invention comprises an isolated or recombinant polynucleotide that encodes a polypeptide that is an ortholog of the isolated polypeptide described above. In one aspect, the invention is an ortholog which, when expressed in *Arabidopsis*, modifies a trait of the *Arabidopsis* plant.

In a fifth aspect, the invention comprises an isolated polypeptide that is a paralog of the isolated polypeptide described above. In one aspect, the invention is a paralog which, when expressed in *Arabidopsis*, modifies a trait of the *Arabidopsis* plant.

In a sixth aspect, the invention comprises an isolated polypeptide that is an ortholog of the isolated polypeptide described above. In one aspect, the invention is an ortholog which, when expressed in *Arabidopsis*, modifies a trait of the *Arabidopsis* plant.

The present invention also encompasses transcription factor variants. A preferred transcription factor variant is one having at least 40% amino acid sequence identity, a more preferred transcription factor variant is one having at least 50% amino acid sequence identity and a most preferred transcription factor variant is one having at least 65% amino acid sequence identity to the transcription factor amino acid sequences SEQ ID NOs: 2 to 2N, where N = 2-123, and which contains at least one functional or structural characteristic of the transcription factor amino acid sequences. Sequences having lesser degrees of identity but comparable biological activity are considered to be equivalents.

In another aspect, the invention is a transgenic plant comprising one or more of the above-described isolated or recombinant polynucleotides. In yet another aspect, the invention is a plant with altered expression levels of a polynucleotide described above or a plant with altered expression or activity levels of an above-described polypeptide. Further, the invention is a plant lacking a nucleotide sequence encoding a polypeptide described above or substantially lacking a polypeptide described above. The plant may be any plant, including, but not limited to, *Arabidopsis*, mustard, soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot,

cauliflower, coffee, cucumber, eggplant, grapes, honeydew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, pumpkin, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits, vegetable brassicas, and mint or other labiates. In yet another aspect, the inventions is an isolated plant material of a plant, including, but not limited to, plant tissue, fruit, seed, plant cell, embryo, protoplast, pollen, and the like. In yet another aspect, the invention is a transgenic plant tissue culture of regenerable cells, including, but not limited to, embryos, meristematic cells, microspores, protoplast, pollen, and the like.

In yet another aspect the invention is a transgenic plant comprising one or more of the above described polynucleotides wherein the encoded polypeptide is expressed and regulates transcription of a gene.

In a further aspect the invention provides a method of using the polynucleotide composition to breed a progeny plant from a transgenic plant including crossing plants, producing seeds from transgenic plants, and methods of breeding using transgenic plants, the method comprising transforming a plant with the polynucleotide composition to create a transgenic plant, crossing the transgenic plant with another plant, selecting seed, and growing the progeny plant from the seed.

In a further aspect, the invention provides a progeny plant derived from a parental plant wherein said progeny plant exhibits at least three fold greater messenger RNA levels than said parental plant, wherein the messenger RNA encodes a DNA-binding protein which is capable of binding to a DNA regulatory sequence and inducing expression of a plant trait gene, wherein the progeny plant is characterized by a change in the plant trait compared to said parental plant. In yet a further aspect, the progeny plant exhibits at least ten fold greater messenger RNA levels compared to said parental plant. In yet a further aspect, the progeny plant exhibits at least fifty fold greater messenger RNA levels compared to said parental plant.

In a further aspect, the invention relates to a cloning or expression vector comprising the isolated or recombinant polynucleotide described above or cells comprising the cloning or expression vector.

In yet a further aspect, the invention relates to a composition produced by incubating a polynucleotide of the invention with a nuclease, a restriction enzyme, a polymerase; a polymerase and a primer; a cloning vector, or with a cell.

Furthermore, the invention relates to a method for producing a plant having a modified trait. The method comprises altering the expression of an isolated or recombinant polynucleotide of the invention or altering the expression or activity of a polypeptide of the invention in a plant to produce a modified plant, and selecting the modified plant for a modified trait. In one aspect, the plant is a monocot plant. In another aspect, the plant is a dicot plant. In another aspect the recombinant polynucleotide is from a dicot plant and the plant is a monocot plant. In yet another aspect the recombinant polynucleotide is from a monocot plant and the plant is a dicot plant. In yet another aspect the recombinant polynucleotide is from a monocot plant and the plant is a monocot plant. In yet another aspect the recombinant polynucleotide is from a dicot plant and the plant is a dicot plant.

In another aspect, the invention is a transgenic plant comprising an isolated or recombinant polynucleotide encoding a polypeptide wherein the polypeptide is selected from the group consisting of SEQ ID NOs: 2-2N where  $N = 2-123$ . In yet another aspect, the invention is a plant with altered expression levels of a polypeptide described above or a plant with altered expression or activity levels of an above-described polypeptide. Further, the invention is a plant lacking a polynucleotide sequence encoding a polypeptide described above or substantially lacking a polypeptide described above. The plant may be any plant, including, but not limited to, Arabidopsis, mustard, soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honeydew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, pumpkin, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits, vegetable brassicas, and mint or other labiates. In yet another aspect, the inventions is an isolated plant material of a plant, including, but not limited to, plant tissue, fruit, seed, plant cell, embryo, protoplast, pollen, and the like. In yet another aspect, the

invention is a transgenic plant tissue culture of regenerable cells, including, but not limited to, embryos, meristematic cells, microspores, protoplast, pollen, and the like.

In another aspect, the invention relates to a method of identifying a factor that is modulated by or interacts with a polypeptide encoded by a polynucleotide of the invention. The method comprises expressing a polypeptide encoded by the polynucleotide in a plant; and identifying at least one factor that is modulated by or interacts with the polypeptide. In one embodiment the method for identifying modulating or interacting factors is by detecting binding by the polypeptide to a promoter sequence, or by detecting interactions between an additional protein and the polypeptide in a yeast two hybrid system, or by detecting expression of a factor by hybridization to a microarray, subtractive hybridization, or differential display.

In yet another aspect, the invention is a method of identifying a molecule that modulates activity or expression of a polynucleotide or polypeptide of interest. The method comprises placing the molecule in contact with a plant comprising the polynucleotide or polypeptide encoded by the polynucleotide of the invention and monitoring one or more of the expression level of the polynucleotide in the plant, the expression level of the polypeptide in the plant, and modulation of an activity of the polypeptide in the plant.

In yet another aspect, the invention relates to an integrated system, computer or computer readable medium comprising one or more character strings corresponding to a polynucleotide of the invention, or to a polypeptide encoded by the polynucleotide. The integrated system, computer or computer readable medium may comprise a link between one or more sequence strings to a modified plant trait.

In yet another aspect, the invention is a method for identifying a sequence similar or homologous to one or more polynucleotides of the invention, or one or more polypeptides encoded by the polynucleotides. The method comprises providing a sequence database, and querying the sequence database with one or more target sequences corresponding to the one or more polynucleotides or to the one or more polypeptides to identify one or more sequence members of the database that display sequence similarity or homology to one or more of the one or more target sequences.

The method may further comprise of linking the one or more of the polynucleotides of the invention, or encoded polypeptides, to a modified plant phenotype.

#### **Brief Description of the Sequence Listing, Tables, and Figure**

The Sequence Listing provides exemplary polynucleotide and polypeptide sequences of the invention. The traits associated with the use of the sequences are included in the Examples.

Diskette1 is a read-only memory computer-readable diskette and contains a copy of the Sequence Listing in ASCII text format. The Sequence Listing is named "SEQLIST514442002040" and is 895 kilobytes in size. The copy of the Sequence Listing on the diskette is hereby incorporated by reference in its entirety.

Table 4 shows the polynucleotides and polypeptides identified by SEQ ID NO; Mendel Gene ID No.; conserved domain of the polypeptide; and if the polynucleotide was tested in a transgenic assay. The first column shows the polynucleotide SEQ ID NO; the second column shows the Mendel Gene ID No., GID; the third column shows the trait(s) resulting from the knock out or overexpression of the polynucleotide in the transgenic plant; the fourth column shows the category of the trait; the fifth column shows the transcription factor family to which the polynucleotide belongs; the sixth column ("Comment"), includes specific effects and utilities conferred by the polynucleotide of the first column; the seventh column shows the SEQ ID NO of the polypeptide encoded by the polynucleotide; and the eighth column shows the amino acid residue positions of the conserved domain in amino acid (AA) co-ordinates.

Table 5 lists a summary of orthologous and homologous sequences identified using BLAST (tblastx program). The first column shows the polynucleotide sequence identifier (SEQ ID NO), the second column shows the corresponding cDNA identifier (Gene ID), the third column shows the orthologous or homologous polynucleotide GenBank Accession Number (Test Sequence ID), the fourth column shows the calculated probability value that the sequence identity is due to chance (Smallest Sum Probability), the fifth column shows the plant species from which the test sequence

was isolated (Test Sequence Species), and the sixth column shows the orthologous or homologous test sequence GenBank annotation (Test Sequence GenBank Annotation).

Figure 1 shows a phylogenic tree of related plant families adapted from Daly et al. (2001 *Plant Physiology* 127:1328-1333).

### **Detailed Description of Exemplary Embodiments**

In an important aspect, the present invention relates to polynucleotides and polypeptides, e.g. for modifying phenotypes of plants. Throughout this disclosure, various information sources are referred to and/or are specifically incorporated. The information sources include scientific journal articles, patent documents, textbooks, and World Wide Web browser-inactive page addresses, for example. While the reference to these information sources clearly indicates that they can be used by one of skill in the art, applicants specifically incorporate each and every one of the information sources cited herein, in their entirety, whether or not a specific mention of "incorporation by reference" is noted. The contents and teachings of each and every one of the information sources can be relied on and used to make and use embodiments of the invention.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a plant" includes a plurality of such plants, and a reference to "a stress" is a reference to one or more stresses and equivalents thereof known to those skilled in the art, and so forth.

The polynucleotide sequences of the invention encode polypeptides that are members of well-known transcription factor families, including plant transcription factor families, as disclosed in Table 4. Generally, the transcription factors encoded by the present sequences are involved in cell differentiation and proliferation and the regulation of growth. Accordingly, one skilled in the art would recognize that by expressing the present sequences in a plant, one may change the expression of autologous genes or induce the expression of introduced genes. By affecting the expression of similar autologous sequences in a plant that have the biological activity

of the present sequences, or by introducing the present sequences into a plant, one may alter a plant's phenotype to one with improved traits. The sequences of the invention may also be used to transform a plant and introduce desirable traits not found in the wild-type cultivar or strain. Plants may then be selected for those that produce the most desirable degree of over- or underexpression of target genes of interest and coincident trait improvement.

The sequences of the present invention may be from any species, particularly plant species, in a naturally occurring form or from any source whether natural, synthetic, semi-synthetic or recombinant. The sequences of the invention may also include fragments of the present amino acid sequences. In this context, a "fragment" refers to a fragment of a polypeptide sequence which is at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity of a transcription factor. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

As one of ordinary skill in the art recognizes, transcription factors can be identified by the presence of a region or domain of structural similarity or identity to a specific consensus sequence or the presence of a specific consensus DNA-binding site or DNA-binding site motif (see, for example, Riechmann et al., (2000) *Science* 290: 2105-2110). The plant transcription factors may belong to one of the following transcription factor families: the AP2 (APETALA2) domain transcription factor family (Riechmann and Meyerowitz (1998) *Biol. Chem.* 379:633-646); the MYB transcription factor family (Martin and Paz-Ares, (1997) *Trends Genet.* 13:67-73); the MADS domain transcription factor family (Riechmann and Meyerowitz (1997) *Biol. Chem.* 378:1079-1101); the WRKY protein family (Ishiguro and Nakamura (1994) *Mol. Gen. Genet.* 244:563-571); the ankyrin-repeat protein family (Zhang et al. (1992) *Plant Cell* 4:1575-1588); the zinc finger protein (Z) family (Klug and Schwabe (1995) *FASEB J.* 9: 597-604); the homeobox (HB) protein family (Buerklin in *Guidebook to the Homeobox Genes*, Duboule (ed.) (1994) Oxford University Press); the CAAT-element binding proteins (Forsburg and Guarente (1989) *Genes Dev.* 3:1166-1178); the squamosa promoter binding proteins (SPB) (Klein et al. (1996)

*Mol. Gen. Genet.* 1996 250:7-16); the NAM protein family (Souer et al. (1996) *Cell* 85:159-170); the IAA/AUX proteins (Rouse et al. (1998) *Science* 279:1371-1373); the HLH/MYC protein family (Littlewood et al. (1994) *Prot. Profile* 1:639-709); the DNA-binding protein (DBP) family (Tucker et al. (1994) *EMBO J.* 13:2994-3002); the bZIP family of transcription factors (Foster et al. (1994) *FASEB J.* 8:192-200); the Box P-binding protein (the BPF-1) family (da Costa e Silva et al. (1993) *Plant J.* 4:125-135); the high mobility group (HMG) family (Bustin and Reeves (1996) *Prog. Nucl. Acids Res. Mol. Biol.* 54:35-100); the scarecrow (SCR) family (Di Laurenzio et al. (1996) *Cell* 86:423-433); the GF14 family (Wu et al. (1997) *Plant Physiol.* 114:1421-1431); the polycomb (PCOMB) family (Kennison (1995) *Annu. Rev. Genet.* 29:289-303); the teosinte branched (TEO) family (Luo et al. (1996) *Nature* 383:794-799); the ABI3 family (Giraudat et al. (1992) *Plant Cell* 4:1251-1261); the triple helix (TH) family (Dehesh et al. (1990) *Science* 250:1397-1399); the EIL family (Chao et al. (1997) *Cell* 89:1133-44); the AT-HOOK family (Reeves and Nissen (1990) *J. Biol. Chem.* 265:8573-8582); the S1FA family (Zhou et al. (1995) *Nucleic Acids Res.* 23:1165-1169); the bZIPT2 family (Lu and Ferl (1995) *Plant Physiol.* 109:723); the YABBY family (Bowman et al. (1999) *Development* 126:2387-96); the PAZ family (Bohmert et al. (1998) *EMBO J.* 17:170-80); a family of miscellaneous (MISC) transcription factors including the DPBF family (Kim et al. (1997) *Plant J.* 11:1237-1251) and the SPF1 family (Ishiguro and Nakamura (1994) *Mol. Gen. Genet.* 244:563-571); the golden (GLD) family (Hall et al. (1998) *Plant Cell* 10:925-936), the TUBBY family (Boggin et al, (1999) *Science* 286:2119-2125), the heat shock family (Wu C (1995) *Annu Rev Cell Dev Biol* 11:441-469), the ENBP family (Christiansen et al (1996) *Plant Mol Biol* 32:809-821), the RING-zinc family (Jensen et al. (1998) *FEBS letters* 436:283-287), the PDBP family (Janik et al *Virology.* (1989) 168:320-329), the PCF family (Cubas P, et al. *Plant J.* (1999) 18:215-22), the SRS (SHI-related) family (Fridborg et al *Plant Cell* (1999) 11:1019-1032), the CPP (cysteine-rich polycomb-like) family (Cvitanich et al *Proc. Natl. Acad. Sci. U S A.* (2000) 97:8163-8168), the ARF (auxin response factor) family (Ulmasov, et al. (1999) *Proc. Natl. Acad. Sci. USA* 96: 5844-5849), the SWI/SNF family (Collingwood et al *J. Mol. End.* 23:255-275), the ACBF family (Seguin et al (1997) *Plant Mol Biol.* 35:281-291), PCGL (CG-1 like) family (da Costa e Silva et al. (1994) *Plant Mol Biol.* 25:921-924) the ARID family (Vazquez et al. (1999) *Development.* 126: 733-42), the Jumonji family, Balciunas et al (2000, *Trends*



*Biochem Sci.* 25: 274-276), the bZIP-NIN family (Schauser et al (1999) *Nature* 402: 191-195), the E2F family Kaelin et al (1992) *Cell* 70: 351-364) and the GRF-like family (Knaap et al (2000) *Plant Physiol.* 122: 695-704). As indicated by any part of the list above and as known in the art, transcription factors have been sometimes categorized by class, family, and sub-family according to their structural content and consensus DNA-binding site motif, for example. Many of the classes and many of the families and sub-families are listed here. However, the inclusion of one sub-family and not another, or the inclusion of one family and not another, does not mean that the invention does not encompass polynucleotides or polypeptides of a certain family or sub-family. The list provided here is merely an example of the types of transcription factors and the knowledge available concerning the consensus sequences and consensus DNA-binding site motifs that help define them as known to those of skill in the art (each of the references noted above are specifically incorporated herein by reference). A transcription factor may include, but is not limited to, any polypeptide that can activate or repress transcription of a single gene or a number of genes. This polypeptide group includes, but is not limited to, DNA-binding proteins, DNA-binding protein binding proteins, protein kinases, protein phosphatases, GTP-binding proteins, and receptors, and the like.

In addition to methods for modifying a plant phenotype by employing one or more polynucleotides and polypeptides of the invention described herein, the polynucleotides and polypeptides of the invention have a variety of additional uses. These uses include their use in the recombinant production (i.e., expression) of proteins; as regulators of plant gene expression, as diagnostic probes for the presence of complementary or partially complementary nucleic acids (including for detection of natural coding nucleic acids); as substrates for further reactions, e.g., mutation reactions, PCR reactions, or the like; as substrates for cloning e.g., including digestion or ligation reactions; and for identifying exogenous or endogenous modulators of the transcription factors. A "polynucleotide" is a nucleic acid sequence comprising a plurality of polymerized nucleotides, e.g., at least about 15 consecutive polymerized nucleotides, optionally at least about 30 consecutive nucleotides, at least about 50 consecutive nucleotides. In many instances, a polynucleotide comprises a nucleotide sequence encoding a polypeptide (or protein) or a domain or fragment thereof. Additionally, the polynucleotide may comprise a promoter, an intron, an enhancer

region, a polyadenylation site, a translation initiation site, 5' or 3' untranslated regions, a reporter gene, a selectable marker, or the like. The polynucleotide can be single stranded or double stranded DNA or RNA. The polynucleotide optionally comprises modified bases or a modified backbone. The polynucleotide can be, e.g., genomic DNA or RNA, a transcript (such as an mRNA), a cDNA, a PCR product, a cloned DNA, a synthetic DNA or RNA, or the like. The polynucleotide can comprise a sequence in either sense or antisense orientations.

A "recombinant polynucleotide" is a polynucleotide that is not in its native state, e.g., the polynucleotide comprises a nucleotide sequence not found in nature, or the polynucleotide is in a context other than that in which it is naturally found, e.g., separated from nucleotide sequences with which it typically is in proximity in nature, or adjacent (or contiguous with) nucleotide sequences with which it typically is not in proximity. For example, the sequence at issue can be cloned into a vector, or otherwise recombined with one or more additional nucleic acid.

An "isolated polynucleotide" is a polynucleotide whether naturally occurring or recombinant, that is present outside the cell in which it is typically found in nature, whether purified or not. Optionally, an isolated polynucleotide is subject to one or more enrichment or purification procedures, e.g., cell lysis, extraction, centrifugation, precipitation, or the like.

A "polypeptide" is an amino acid sequence comprising a plurality of consecutive polymerized amino acid residues e.g., at least about 15 consecutive polymerized amino acid residues, optionally at least about 30 consecutive polymerized amino acid residues, at least about 50 consecutive polymerized amino acid residues. In many instances, a polypeptide comprises a polymerized amino acid residue sequence that is a transcription factor or a domain or portion or fragment thereof. Additionally, the polypeptide may comprise a localization domain, 2) an activation domain, 3) a repression domain, 4) an oligomerization domain or 5) a DNA-binding domain, or the like. The polypeptide optionally comprises modified amino acid residues, naturally occurring amino acid residues not encoded by a codon, non-naturally occurring amino acid residues.

A "recombinant polypeptide" is a polypeptide produced by translation of a recombinant polynucleotide. A "synthetic polypeptide" is a polypeptide created by consecutive polymerization of isolated amino acid residues using methods well known in the art. An "isolated polypeptide," whether a naturally occurring or a recombinant polypeptide, is more enriched in (or out of) a cell than the polypeptide in its natural state in a wild type cell, e.g., more than about 5% enriched, more than about 10% enriched, or more than about 20%, or more than about 50%, or more, enriched, i.e., alternatively denoted: 105%, 110%, 120%, 150% or more, enriched relative to wild type standardized at 100%. Such an enrichment is not the result of a natural response of a wild type plant. Alternatively, or additionally, the isolated polypeptide is separated from other cellular components with which it is typically associated, e.g., by any of the various protein purification methods herein.

"Identity" or "similarity" refers to sequence similarity between two polynucleotide sequences or between two polypeptide sequences, with identity being a more strict comparison. The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more polynucleotide sequences or two or more polypeptide sequences. Identity or similarity can be determined by comparing a position in each sequence that may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same nucleotide base or amino acid, then the molecules are identical at that position. A degree of similarity or identity between polynucleotide sequences is a function of the number of identical or matching nucleotides at positions shared by the polynucleotide sequences. A degree of identity of polypeptide sequences is a function of the number of identical amino acids at positions shared by the polypeptide sequences. A degree of homology or similarity of polypeptide sequences is a function of the number of amino acids, i.e., structurally related, at positions shared by the polypeptide sequences.

"Altered" nucleic acid sequences encoding polypeptide include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide encoding a polypeptide with at least one functional characteristic of the polypeptide. Included within this definition are polymorphisms that may or may not be readily detectable using a particular oligonucleotide probe of the

polynucleotide encoding polypeptide, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding polypeptide. The encoded polypeptide protein may also be "altered", and may contain deletions, insertions, or substitutions of amino acid residues that produce a silent change and result in a functionally equivalent polypeptide. Deliberate amino acid substitutions may be made on the basis of similarity in residue side chain chemistry, including, but not limited to, polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological activity of polypeptide is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine. Alignments between different polypeptide sequences may be used to calculate "percentage sequence similarity".

The term "plant" includes whole plants, shoot vegetative organs/structures (*e.g.*, leaves, stems and tubers), roots, flowers and floral organs/structures (*e.g.*, bracts, sepals, petals, stamens, carpels, anthers and ovules), seed (including embryo, endosperm, and seed coat) and fruit (the mature ovary), plant tissue (*e.g.*, vascular tissue, ground tissue, and the like) and cells (*e.g.*, guard cells, egg cells, and the like), and progeny of same. The class of plants that can be used in the method of the invention is generally as broad as the class of higher and lower plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), gymnosperms, ferns, horsetails, psilophytes, lycophytes, bryophytes, and multicellular algae. (See for example, Figure 1, adapted from Daly et al. 2001 *Plant Physiology* 127:1328-1333; and see also Tudge, C., The Variety of Life, Oxford University Press, New York, 2000, pp. 547-606.)

A "transgenic plant" refers to a plant that contains genetic material not found in a wild type plant of the same species, variety or cultivar. The genetic material may include a transgene, an insertional mutagenesis event (such as by transposon or T-DNA insertional mutagenesis), an activation tagging sequence, a mutated sequence, a homologous recombination event or a sequence modified by chimera-plasty.

Typically, the foreign genetic material has been introduced into the plant by human manipulation, but any method can be used as one of skill in the art recognizes.

A transgenic plant may contain an expression vector or cassette. The expression cassette typically comprises a polypeptide-encoding sequence operably linked (i.e., under regulatory control of) to appropriate inducible or constitutive regulatory sequences that allow for the expression of polypeptide. The expression cassette can be introduced into a plant by transformation or by breeding after transformation of a parent plant. A plant refers to a whole plant as well as to a plant part, such as seed, fruit, leaf, or root, plant tissue, plant cells or any other plant material, e.g., a plant explant, as well as to progeny thereof, and to *in vitro* systems that mimic biochemical or cellular components or processes in a cell.

"Ectopic expression or altered expression" in reference to a polynucleotide indicates that the pattern of expression in, e.g., a transgenic plant or plant tissue, is different from the expression pattern in a wild type plant or a reference plant of the same species. The pattern of expression may also be compared with a reference expression pattern in a wild type plant of the same species. For example, the polynucleotide or polypeptide is expressed in a cell or tissue type other than a cell or tissue type in which the sequence is expressed in the wild type plant, or by expression at a time other than at the time the sequence is expressed in the wild type plant, or by a response to different inducible agents, such as hormones or environmental signals, or at different expression levels (either higher or lower) compared with those found in a wild type plant. The term also refers to altered expression patterns that are produced by lowering the levels of expression to below the detection level or completely abolishing expression. The resulting expression pattern can be transient or stable, constitutive or inducible. In reference to a polypeptide, the term "ectopic expression or altered expression" further may relate to altered activity levels resulting from the interactions of the polypeptides with exogenous or endogenous modulators or from interactions with factors or as a result of the chemical modification of the polypeptides.

A "fragment" or "domain," with respect to a polypeptide, refers to a subsequence of the polypeptide. In some cases, the fragment or domain, is a subsequence of the polypeptide which performs at least one biological function of the

intact polypeptide in substantially the same manner, or to a similar extent, as does the intact polypeptide. For example, a polypeptide fragment can comprise a recognizable structural motif or functional domain such as a DNA-binding site or domain that binds to a DNA promoter region, an activation domain, or a domain for protein-protein interactions. Fragments can vary in size from as few as 6 amino acids to the full length of the intact polypeptide, but are preferably at least about 30 amino acids in length and more preferably at least about 60 amino acids in length. In reference to a polynucleotide sequence, "a fragment" refers to any subsequence of a polynucleotide, typically, of at least about 15 consecutive nucleotides, preferably at least about 30 nucleotides, more preferably at least about 50 nucleotides, of any of the sequences provided herein.

The invention also encompasses production of DNA sequences that encode transcription factors and transcription factor derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding transcription factors or any fragment thereof.

A "conserved domain", with respect to a polypeptide, refers to a domain within a transcription factor family which exhibits a higher degree of sequence homology, such as at least 65% sequence identity including conservative substitutions, and preferably at least 80% sequence identity, and more preferably at least 85%, or at least about 86%, or at least about 87%, or at least about 88%, or at least about 90%, or at least about 95%, or at least about 98% amino acid residue sequence identity of a polypeptide of consecutive amino acid residues. A fragment or domain can be referred to as outside a consensus sequence or outside a consensus DNA-binding site that is known to exist or that exists for a particular transcription factor class, family, or sub-family. In this case, the fragment or domain will not include the exact amino acids of a consensus sequence or consensus DNA-binding site of a transcription factor class, family or sub-family, or the exact amino acids of a particular transcription factor consensus sequence or consensus DNA-binding site. Furthermore, a particular fragment, region, or domain of a polypeptide, or a polynucleotide encoding a polypeptide, can be "outside a conserved domain" if all the amino acids of the

fragment, region, or domain fall outside of a defined conserved domain(s) for a polypeptide or protein. The conserved domains for each of polypeptides of SEQ ID NOs:2 - 2N, where N = 2-123, are listed in Table 4 as described in Example VII. Also, many of the polypeptides of Table 4 have conserved domains specifically indicated by start and stop sites. A comparison of the regions of the polypeptides in SEQ ID NOs:2 - 2N, where N = 2-123, or of those in Table 4, allows one of skill in the art to identify conserved domain(s) for any of the polypeptides listed or referred to in this disclosure, including those in Table 4 or Table 5.

A "trait" refers to a physiological, morphological, biochemical, or physical characteristic of a plant or particular plant material or cell. In some instances, this characteristic is visible to the human eye, such as seed or plant size, or can be measured by biochemical techniques, such as detecting the protein, starch, or oil content of seed or leaves, or by observation of a metabolic or physiological process, e.g. by measuring uptake of carbon dioxide, or by the observation of the expression level of a gene or genes, e.g., by employing Northern analysis, RT-PCR, microarray gene expression assays, or reporter gene expression systems, or by agricultural observations such as stress tolerance, yield, or pathogen tolerance. Any technique can be used to measure the amount of, comparative level of, or difference in any selected chemical compound or macromolecule in the transgenic plants, however.

"Trait modification" refers to a detectable difference in a characteristic in a plant ectopically expressing a polynucleotide or polypeptide of the present invention relative to a plant not doing so, such as a wild type plant. In some cases, the trait modification can be evaluated quantitatively. For example, the trait modification can entail at least about a 2% increase or decrease in an observed trait (difference), at least a 5% difference, at least about a 10% difference, at least about a 20% difference, at least about a 30%, at least about a 50%, at least about a 70%, or at least about a 100%, or an even greater difference compared with a wild type plant. It is known that there can be a natural variation in the modified trait. Therefore, the trait modification observed entails a change of the normal distribution of the trait in the plants compared with the distribution observed in wild type plant.

## **I. Traits Which May Be Modified**

Trait modifications of particular interest include those to seed (such as embryo or endosperm), fruit, root, flower, leaf, stem, shoot, seedling or the like, including: enhanced tolerance to environmental conditions including freezing, chilling, heat, drought, water saturation, radiation and ozone; improved tolerance to microbial, fungal or viral diseases; improved tolerance to pest infestations, including nematodes, mollicutes, parasitic higher plants or the like; decreased herbicide sensitivity; improved tolerance of heavy metals or enhanced ability to take up heavy metals; improved growth under poor photoconditions (e.g., low light and/or short day length), or changes in expression levels of genes of interest. Other phenotype that can be modified relate to the production of plant metabolites, such as variations in the production of taxol, tocopherol, tocotrienol, sterols, phytosterols, vitamins, wax monomers, anti-oxidants, amino acids, lignins, cellulose, tannins, prenolipids (such as chlorophylls and carotenoids), glucosinolates, and terpenoids, enhanced or compositionally altered protein or oil production (especially in seeds), or modified sugar (insoluble or soluble) and/or starch composition. Physical plant characteristics that can be modified include cell development (such as the number of trichomes), fruit and seed size and number, yields of plant parts such as stems, leaves, inflorescences, and roots, the stability of the seeds during storage, characteristics of the seed pod (e.g., susceptibility to shattering), root hair length and quantity, internode distances, or the quality of seed coat. Plant growth characteristics that can be modified include growth rate, germination rate of seeds, vigor of plants and seedlings, leaf and flower senescence, male sterility, apomixis, flowering time, flower abscission, rate of nitrogen uptake, osmotic sensitivity to soluble sugar concentrations, biomass or transpiration characteristics, as well as plant architecture characteristics such as apical dominance, branching patterns, number of organs, organ identity, organ shape or size.

## **II. Transcription Factors Modify Expression Of Endogenous Genes**

Expression of genes which encode transcription factors modify expression of endogenous genes, polynucleotides, and proteins are well known in the art. In addition, transgenic plants comprising isolated polynucleotides encoding transcription factors may also modify expression of endogenous genes, polynucleotides, and proteins. Examples include Peng et al. (1997, *Genes and Development* 11:3194-3205) and Peng et al. (1999, *Nature*, 400:256-261). In addition, many others have



demonstrated that an *Arabidopsis* transcription factor expressed in an exogenous plant species elicits the same or very similar phenotypic response. See, for example, Fu et al. (2001, *Plant Cell* 13:1791-1802); Nandi et al. (2000, *Curr. Biol.* 10:215-218); Coupland (1995, *Nature* 377:482-483); and Weigel and Nilsson (1995, *Nature* 377:482-500).

In another example, Mandel et al. (1992, *Cell* 71:133-143) and Suzuki et al. (2001, *Plant J.* 28:409-418) teach that a transcription factor expressed in another plant species elicits the same or very similar phenotypic response of the endogenous sequence, as often predicted in earlier studies of *Arabidopsis* transcription factors in *Arabidopsis* (see Mandel et al., 1992, *supra*; Suzuki et al., 2001, *supra*).

Other examples include Müller et al. (2001, *Plant J.* 28:169-179); Kim et al. (2001, *Plant J.* 25:247-259); Kyoizuka and Shimamoto (2002, *Plant Cell Physiol.* 43:130-135); Boss and Thomas (2002, *Nature*, 416:847-850); He et al. (2000, *Transgenic Res.*, 9:223-227); and Robson et al. (2001, *Plant J.* 28:619-631).

In yet another example, Gilmour et al. (1998, *Plant J.* 16:433-442) teach an *Arabidopsis* AP2 transcription factor, CBF1, which, when overexpressed in transgenic plants, increases plant freezing tolerance. Jaglo et al (2001, *Plant Physiol.* 127:910-917) further identified sequences in *Brassica napus* which encode CBF-like genes and that transcripts for these genes accumulated rapidly in response to low temperature. Transcripts encoding CBF-like proteins were also found to accumulate rapidly in response to low temperature in wheat, as well as in tomato. An alignment of the CBF proteins from *Arabidopsis*, *B. napus*, wheat, rye, and tomato revealed the presence of conserved amino acid sequences, PKK/RPAGR×KF×ETRHP and DSAWR, that bracket the AP2/EREBP DNA binding domains of the proteins and distinguish them from other members of the AP2/EREBP protein family. (See Jaglo et al., *supra*.)

### III. Polypeptides and Polynucleotides of the Invention

The present invention provides, among other things, transcription factors (TFs), and transcription factor homologue polypeptides, and isolated or recombinant polynucleotides encoding the polypeptides, or novel variant polypeptides or polynucleotides encoding novel variants of transcription factors derived from the specific sequences provided here. These polypeptides and polynucleotides may be employed to modify a plant's characteristic.

Exemplary polynucleotides encoding the polypeptides of the invention were identified in the *Arabidopsis thaliana* GenBank database using publicly available sequence analysis programs and parameters. Sequences initially identified were then further characterized to identify sequences comprising specified sequence strings corresponding to sequence motifs present in families of known transcription factors. In addition, further exemplary polynucleotides encoding the polypeptides of the invention were identified in the plant GenBank database using publicly available sequence analysis programs and parameters. Sequences initially identified were then further characterized to identify sequences comprising specified sequence strings corresponding to sequence motifs present in families of known transcription factors. Polynucleotide sequences meeting such criteria were confirmed as transcription factors.

Additional polynucleotides of the invention were identified by screening *Arabidopsis thaliana* and/or other plant cDNA libraries with probes corresponding to known transcription factors under low stringency hybridization conditions. Additional sequences, including full length coding sequences were subsequently recovered by the rapid amplification of cDNA ends (RACE) procedure, using a commercially available kit according to the manufacturer's instructions. Where necessary, multiple rounds of RACE are performed to isolate 5' and 3' ends. The full length cDNA was then recovered by a routine end-to-end polymerase chain reaction (PCR) using primers specific to the isolated 5' and 3' ends. Exemplary sequences are provided in the Sequence Listing.

The polynucleotides of the invention can be or were ectopically expressed in overexpressor or knockout plants and the changes in the characteristic(s) or trait(s) of the plants observed. Therefore, the polynucleotides and polypeptides can be employed to improve the characteristics of plants.

The polynucleotides of the invention can be or were ectopically expressed in overexpressor plant cells and the changes in the expression levels of a number of genes, polynucleotides, and/or proteins of the plant cells observed. Therefore, the polynucleotides and polypeptides can be employed to change expression levels of a genes, polynucleotides, and/or proteins of plants.

#### **IV. Producing Polypeptides**

The polynucleotides of the invention include sequences that encode transcription factors and transcription factor homologue polypeptides and sequences complementary thereto, as well as unique fragments of coding sequence, or sequence complementary thereto. Such polynucleotides can be, e.g., DNA or RNA, e.g., mRNA, cRNA, synthetic RNA, genomic DNA, cDNA synthetic DNA, oligonucleotides, etc. The polynucleotides are either double-stranded or single-stranded, and include either, or both sense (i.e., coding) sequences and antisense (i.e., non-coding, complementary) sequences. The polynucleotides include the coding sequence of a transcription factor, or transcription factor homologue polypeptide, in isolation, in combination with additional coding sequences (e.g., a purification tag, a localization signal, as a fusion-protein, as a pre-protein, or the like), in combination with non-coding sequences (e.g., introns or inteins, regulatory elements such as promoters, enhancers, terminators, and the like), and/or in a vector or host environment in which the polynucleotide encoding a transcription factor or transcription factor homologue polypeptide is an endogenous or exogenous gene.

A variety of methods exist for producing the polynucleotides of the invention. Procedures for identifying and isolating DNA clones are well known to those of skill in the art, and are described in, e.g., Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA ("Berger"); Sambrook et al., Molecular Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989

("Sambrook") and Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2000) ("Ausubel").

Alternatively, polynucleotides of the invention, can be produced by a variety of in vitro amplification methods adapted to the present invention by appropriate selection of specific or degenerate primers. Examples of protocols sufficient to direct persons of skill through in vitro amplification methods, including the polymerase chain reaction (PCR) the ligase chain reaction (LCR), Qbeta-replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA), e.g., for the production of the homologous nucleic acids of the invention are found in Berger (*supra*), Sambrook (*supra*), and Ausubel (*supra*), as well as Mullis et al., (1987) PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press Inc. San Diego, CA (1990) (Innis). Improved methods for cloning in vitro amplified nucleic acids are described in Wallace et al., U.S. Pat. No. 5,426,039. Improved methods for amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369: 684-685 and the references cited therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See, e.g., Ausubel, Sambrook and Berger, all supra.*

Alternatively, polynucleotides and oligonucleotides of the invention can be assembled from fragments produced by solid-phase synthesis methods. Typically, fragments of up to approximately 100 bases are individually synthesized and then enzymatically or chemically ligated to produce a desired sequence, e.g., a polynucleotide encoding all or part of a transcription factor. For example, chemical synthesis using the phosphoramidite method is described, e.g., by Beaucage et al. (1981) Tetrahedron Letters 22:1859-1869; and Matthes et al. (1984) EMBO J. 3:801-805. According to such methods, oligonucleotides are synthesized, purified, annealed to their complementary strand, ligated and then optionally cloned into suitable vectors. And if so desired, the polynucleotides and polypeptides of the invention can be custom ordered from any of a number of commercial suppliers.

## V. Homologous Sequences

Sequences homologous, i.e., that share significant sequence identity or similarity, to those provided in the Sequence Listing, derived from *Arabidopsis thaliana* or from other plants of choice are also an aspect of the invention.

Homologous sequences can be derived from any plant including monocots and dicots and in particular agriculturally important plant species, including but not limited to, crops such as soybean, wheat, corn, potato, cotton, rice, rape, oilseed rape (including canola), sunflower, alfalfa, sugarcane and turf; or fruits and vegetables, such as banana, blackberry, blueberry, strawberry, and raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honeydew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, pumpkin, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits (such as apple, peach, pear, cherry and plum) and vegetable brassicas (such as broccoli, cabbage, cauliflower, Brussels sprouts, and kohlrabi). Other crops, fruits and vegetables whose phenotype can be changed include barley, rye, millet, sorghum, currant, avocado, citrus fruits such as oranges, lemons, grapefruit and tangerines, artichoke, cherries, nuts such as the walnut and peanut, endive, leek, roots, such as arrowroot, beet, cassava, turnip, radish, yam, and sweet potato, and beans. The homologous sequences may also be derived from woody species, such pine, poplar and eucalyptus, or mint or other labiates.

### Orthologs And Paralogs

Several different methods are known by those of skill in the art for identifying and defining these functionally homologous sequences. Three general methods for defining paralogs and orthologs are described; a paralog or ortholog or homolog may be identified by one or more of the methods described below.

Orthologs and paralogs are evolutionarily related genes that have similar sequence and similar functions. Orthologs are structurally related genes in different species most likely derived from a common ancestral molecule prior to speciation. Paralogs are structurally related genes within a single species and are most likely a result of gene and/or chromosomal duplication.

Within a single plant species, gene duplication may causes two copies of a particular gene, giving rise to two or more genes with similar sequence and similar

function known as paralogs. A paralog is therefore a similar gene with a similar function within the same species. Paralogs typically cluster together or in the same clade (a group of similar genes) when a gene family phylogeny is analyzed using programs such as CLUSTAL (Thompson et al. (1994) *Nucleic Acids Res.* 22:4673-4680; Higgins et al. (1996) *Methods Enzymol.* 266 383-402). Groups of similar genes can also be identified using by pair-wise BLAST analysis (Feng and Doolittle (1987) *J. Mol. Evol.* 25:351-360). For example, a clade of very similar MADS domain transcription factors from *Arabidopsis* all share a common function in flowering time (Ratcliffe et al. (2001) *Plant Physiol.* 126:122-132), and a group of very similar AP2 domain transcription factors from *Arabidopsis* are involved in tolerance of plants to freezing (Gilmour et al. (1998) *Plant J.* 16:433-442). Analysis of groups of similar genes with similar function that fall within one clade can yield sub-sequences that are particular to the clade. These sub-sequences, known as consensus sequences, can not only be used to define the sequences within each clade, but define the functions of these genes, since genes within each clade typically share the same function. (See also, for example, Mount, D.W. (2001) Bioinformatics: Sequence and Genome Analysis Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York page 543.)

Speciation, the production of new species from a parental species, can also give rise to two or more genes with similar sequence and similar function. These genes, termed orthologs, often have an identical function within their host plants and are often interchangeable between species without losing function. Because plants have common ancestors, many genes in any plant species will have a corresponding orthologous gene in another plant species. Once a phylogenic tree for a gene family of one species has been constructed using a program such as CLUSTAL (Thompson et al. (1994) *Nucleic Acids Res.* 22:4673-4680; Higgins et al. (1996) *Methods Enzymol.* 266:383-402), potential orthologous sequences can be placed into the phylogenetic tree and its relationship to genes from the species of interest can be determined. Once the ortholog pair has been identified, the function of the test ortholog can be determined by determining the function of the reference ortholog.

Transcription factors that are homologous to the listed sequences will typically share at least about 30% amino acid sequence identity, or at least about 30% amino

acid sequence identity outside of a known consensus sequence or consensus DNA-binding site. More closely related transcription factors can share at least about 50%, about 60%, about 65%, about 70%, about 75% or about 80% or about 90% or about 95% or about 98% or more sequence identity with the listed sequences, or with the listed sequences but excluding or outside a known consensus sequence or consensus DNA-binding site, or with the listed sequences excluding one or all conserved domain. Factors that are most closely related to the listed sequences share, e.g., at least about 85%, about 90% or about 95% or more % sequence identity to the listed sequences, or to the listed sequences but excluding or outside a known consensus sequence or consensus DNA-binding site or outside one or all conserved domain. At the nucleotide level, the sequences will typically share at least about 40% nucleotide sequence identity, preferably at least about 50%, about 60%, about 70% or about 80% sequence identity, and more preferably about 85%, about 90%, about 95% or about 97% or more sequence identity to one or more of the listed sequences, or to a listed sequence but excluding or outside a known consensus sequence or consensus DNA-binding site, or outside one or all conserved domain. The degeneracy of the genetic code enables major variations in the nucleotide sequence of a polynucleotide while maintaining the amino acid sequence of the encoded protein. Conserved domains within a transcription factor family may exhibit a higher degree of sequence homology, such as at least 65% sequence identity including conservative substitutions, and preferably at least 80% sequence identity, and more preferably at least 85%, or at least about 86%, or at least about 87%, or at least about 88%, or at least about 90%, or at least about 95%, or at least about 98% sequence identity. Transcription factors that are homologous to the listed sequences should share at least 30%, or at least about 60%, or at least about 75%, or at least about 80%, or at least about 90%, or at least about 95% amino acid sequence identity over the entire length of the polypeptide or the homolog. In addition, transcription factors that are homologous to the listed sequences should share at least 30%, or at least about 60%, or at least about 75%, or at least about 80%, or at least about 90%, or at least about 95% amino acid sequence similarity over the entire length of the polypeptide or the homolog.

Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Inc. Madison, Wis.). The MEGALIGN program

can create alignments between two or more sequences according to different methods, e.g., the clustal method. (See, e.g., Higgins, D. G. and P. M. Sharp (1988) *Gene* 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. Other alignment algorithms or programs may be used, including FASTA, BLAST, or ENTREZ, FASTA and BLAST. These are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.), and can be used with or without default settings. ENTREZ is available through the National Center for Biotechnology Information. In one embodiment, the percent identity of two sequences can be determined by the GCG program with a gap weight of 1, e.g., each amino acid gap is weighted as if it were a single amino acid or nucleotide mismatch between the two sequences (see USPN 6,262,333).

Other techniques for alignment are described in *Methods in Enzymology*, vol. 266: *Computer Methods for Macromolecular Sequence Analysis* (1996), ed. Doolittle, Academic Press, Inc., San Diego, Calif., USA. Preferably, an alignment program that permits gaps in the sequence is utilized to align the sequences. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See *Methods Mol. Biol.* 70: 173-187 (1997). Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. An alternative search strategy uses MPSRCH software, which runs on a MASPAC computer. MPSRCH uses a Smith-Waterman algorithm to score sequences on a massively parallel computer. This approach improves ability to pick up distantly related matches, and is especially tolerant of small gaps and nucleotide sequence errors. Nucleic acid-encoded amino acid sequences can be used to search both protein and DNA databases.

The percentage similarity between two polypeptide sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between polynucleotide sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) *Methods Enzymol.*



183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions (see US Patent Application No. 20010010913).

Thus, the invention provides methods for identifying a sequence similar or paralogous or orthologous or homologous to one or more polynucleotides as noted herein, or one or more target polypeptides encoded by the polynucleotides, or otherwise noted herein and may include linking or associating a given plant phenotype or gene function with a sequence. In the methods, a sequence database is provided (locally or across an inter or intra net) and a query is made against the sequence database using the relevant sequences herein and associated plant phenotypes or gene functions.

In addition, one or more polynucleotide sequences or one or more polypeptides encoded by the polynucleotide sequences may be used to search against a BLOCKS (Bairoch et al. (1997) *Nucleic Acids Res.* 25:217-221), PFAM, and other databases which contain previously identified and annotated motifs, sequences and gene functions. Methods that search for primary sequence patterns with secondary structure gap penalties (Smith et al. (1992) *Protein Engineering* 5:35-51) as well as algorithms such as Basic Local Alignment Search Tool (BLAST; Altschul, S. F. (1993) *J. Mol. Evol.* 36:290-300; Altschul et al. (1990) *supra*), BLOCKS (Henikoff, S. and Henikoff, G. J. (1991) *Nucleic Acids Research* 19:6565-6572), Hidden Markov Models (HMM; Eddy, S. R. (1996) *Cur. Opin. Str. Biol.* 6:361-365; Sonnhammer et al. (1997) *Proteins* 28:405-420), and the like, can be used to manipulate and analyze polynucleotide and polypeptide sequences encoded by polynucleotides. These databases, algorithms and other methods are well known in the art and are described in Ausubel et al. (1997; *Short Protocols in Molecular Biology*, John Wiley & Sons, New York N.Y., unit 7.7) and in Meyers, R. A. (1995; *Molecular Biology and Biotechnology*, Wiley VCH, New York N.Y., p 856-853).

Furthermore, methods using manual alignment of sequences similar or homologous to one or more polynucleotide sequences or one or more polypeptides encoded by the polynucleotide sequences may be used to identify regions of similarity and conserved domains. Such manual methods are well-known of those of skill in the

art and can include, for example, comparisons of tertiary structure between a polypeptide sequence encoded by a polynucleotide which comprises a known function with a polypeptide sequence encoded by a polynucleotide sequence which has a function not yet determined. Such examples of tertiary structure may comprise predicted alpha helices, beta-sheets, amphipathic helices, leucine zipper motifs, zinc finger motifs, proline-rich regions, cysteine repeat motifs, and the like.

## **VI. Identifying Polynucleotides or Nucleic Acids by Hybridization**

Polynucleotides homologous to the sequences illustrated in the Sequence Listing and tables can be identified, e.g., by hybridization to each other under stringent or under highly stringent conditions. Single stranded polynucleotides hybridize when they associate based on a variety of well characterized physical-chemical forces, such as hydrogen bonding, solvent exclusion, base stacking and the like. The stringency of a hybridization reflects the degree of sequence identity of the nucleic acids involved, such that the higher the stringency, the more similar are the two polynucleotide strands. Stringency is influenced by a variety of factors, including temperature, salt concentration and composition, organic and non-organic additives, solvents, etc. present in both the hybridization and wash solutions and incubations (and number thereof), as described in more detail in the references cited above. Encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NOS: 26; 46; 176; 114; 142; 144; 82; 50; 72; 96; 18; 22; 26; and 240, and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G. M. and S. L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A. R. (1987) *Methods Enzymol.* 152:507-511.) Estimates of homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds. (1985) *Nucleic Acid Hybridisation*, IRL Press, Oxford, U.K.). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions.

In addition to the nucleotide sequences listed in Tables 4 and 5, full length cDNA, orthologs, paralogs and homologs of the present nucleotide sequences may be identified and isolated using well known methods. The cDNA libraries orthologs, paralogs and homologs of the present nucleotide sequences may be screened using hybridization methods to determine their utility as hybridization target or amplification probes.

An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Nucleic acid molecules that hybridize under stringent conditions will typically hybridize to a probe based on either the entire cDNA or selected portions, e.g., to a unique subsequence, of the cDNA under wash conditions of 0.2x SSC to 2.0 x SSC, 0.1% SDS at 50-65° C. For example, high stringency is about 0.2 x SSC, 0.1% SDS at 65° C. Ultra-high stringency will be the same conditions except the wash temperature is raised about 3 to about 5° C, and ultra-ultra-high stringency will be the same conditions except the wash temperature is raised about 6 to about 9° C. For identification of less closely related homologues washes can be performed at a lower temperature, e.g., 50° C. In general, stringency is increased by raising the wash temperature and/or decreasing the concentration of SSC, as known in the art.

In another example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30° C, more preferably of at least about 37° C, and most preferably of at least about 42° C. Varying additional parameters, such as

hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30° C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37° C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 µg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42° C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50% formamide, and 200 µg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps that follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25° C, more preferably of at least about 42° C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65° C. The most preferred high stringency washes are of at least about 68° C. For example, in a preferred embodiment, wash steps will occur at 25° C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42° C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, the wash steps will occur at 68° C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art (see U.S. Patent Application No. 20010010913).

As another example, stringent conditions can be selected such that an oligonucleotide that is perfectly complementary to the coding oligonucleotide hybridizes to the coding oligonucleotide with at least about a 5-10x higher signal to

noise ratio than the ratio for hybridization of the perfectly complementary oligonucleotide to a nucleic acid encoding a transcription factor known as of the filing date of the application. Conditions can be selected such that a higher signal to noise ratio is observed in the particular assay which is used, e.g., about 15x, 25x, 35x, 50x or more. Accordingly, the subject nucleic acid hybridizes to the unique coding oligonucleotide with at least a 2x higher signal to noise ratio as compared to hybridization of the coding oligonucleotide to a nucleic acid encoding known polypeptide. Again, higher signal to noise ratios can be selected, e.g., about 5x, 10x, 25x, 35x, 50x or more. The particular signal will depend on the label used in the relevant assay, e.g., a fluorescent label, a colorimetric label, a radioactive label, or the like.

Alternatively, transcription factor homolog polypeptides can be obtained by screening an expression library using antibodies specific for one or more transcription factors. With the provision herein of the disclosed transcription factor, and transcription factor homologue nucleic acid sequences, the encoded polypeptide(s) can be expressed and purified in a heterologous expression system (e.g., *E. coli*) and used to raise antibodies (monoclonal or polyclonal) specific for the polypeptide(s) in question. Antibodies can also be raised against synthetic peptides derived from transcription factor, or transcription factor homologue, amino acid sequences. Methods of raising antibodies are well known in the art and are described in Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, New York. Such antibodies can then be used to screen an expression library produced from the plant from which it is desired to clone additional transcription factor homologues, using the methods described above. The selected cDNAs can be confirmed by sequencing and enzymatic activity.

## VII. Sequence Variations

It will readily be appreciated by those of skill in the art, that any of a variety of polynucleotide sequences are capable of encoding the transcription factors and transcription factor homologue polypeptides of the invention. Due to the degeneracy of the genetic code, many different polynucleotides can encode identical and/or substantially similar polypeptides in addition to those sequences illustrated in the Sequence Listing. Nucleic acids having a sequence that differs from the sequences

shown in the Sequence Listing, or complementary sequences, that encode functionally equivalent peptides (i.e., peptides having some degree of equivalent or similar biological activity) but differ in sequence from the sequence shown in the sequence listing due to degeneracy in the genetic code, are also within the scope of the invention.

Altered polynucleotide sequences encoding polypeptides include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide encoding a polypeptide with at least one functional characteristic of the instant polypeptides. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding the instant polypeptides, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding the instant polypeptides.

Allelic variant refers to any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations can be silent (i.e., no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequence. The term allelic variant is also used herein to denote a protein encoded by an allelic variant of a gene. Splice variant refers to alternative forms of RNA transcribed from a gene. Splice variation arises naturally through use of alternative splicing sites within a transcribed RNA molecule, or less commonly between separately transcribed RNA molecules, and may result in several mRNAs transcribed from the same gene. Splice variants may encode polypeptides having altered amino acid sequence. The term splice variant is also used herein to denote a protein encoded by a splice variant of an mRNA transcribed from a gene.

Those skilled in the art would recognize that the polypeptide sequence G481, SEQ ID NO: 114, represents a single transcription factor; allelic variation and alternative splicing may be expected to occur. Allelic variants of the polynucleotide sequence of SEQ ID NO: 113 can be cloned by probing cDNA or genomic libraries from different individual organisms according to standard procedures. Allelic variants of the DNA sequence shown in SEQ ID NO:113, including those containing

silent mutations and those in which mutations result in amino acid sequence changes, are within the scope of the present invention, as are proteins which are allelic variants of SEQ ID NO: 114. cDNAs generated from alternatively spliced mRNAs, which retain the properties of the transcription factor are included within the scope of the present invention, as are polypeptides encoded by such cDNAs and mRNAs. Allelic variants and splice variants of these sequences can be cloned by probing cDNA or genomic libraries from different individual organisms or tissues according to standard procedures known in the art (see USPN 6,388,064).

For example, Table 1 illustrates, e.g., that the codons AGC, AGT, TCA, TCC, TCG, and TCT all encode the same amino acid: serine. Accordingly, at each position in the sequence where there is a codon encoding serine, any of the above trinucleotide sequences can be used without altering the encoded polypeptide.

**Table 1**

Amino acid			Possible Codons							
Alanine	Ala	A	GCA	GCC	GCG	GCU				
Cysteine	Cys	C	TGC	TGT						
Aspartic acid	Asp	D	GAC	GAT						
Glutamic acid	Glu	E	GAA	GAG						
Phenylalanine	Phe	F	TTC	TTT						
Glycine	Gly	G	GGA	GGC	GGG	GGT				
Histidine	His	H	CAC	CAT						
Isoleucine	Ile	I	ATA	ATC	ATT					
Lysine	Lys	K	AAA	AAG						
Leucine	Leu	L	TTA	TTG	CTA	CTC	CTG	CTT		
Methionine	Met	M	ATG							
Asparagine	Asn	N	AAC	AAT						
Proline	Pro	P	CCA	CCC	CCG	CCT				
Glutamine	Gln	Q	CAA	CAG						
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGT		
Serine	Ser	S	AGC	AGT	TCA	TCC	TCG	TCT		
Threonine	Thr	T	ACA	ACC	ACG	ACT				
Valine	Val	V	GTA	GTC	GTG	GTT				
Tryptophan	Trp	W	TGG							
Tyrosine	Tyr	Y	TAC	TAT						

Sequence alterations that do not change the amino acid sequence encoded by the polynucleotide are termed "silent" variations. With the exception of the codons ATG

and TGG, encoding methionine and tryptophan, respectively, any of the possible codons for the same amino acid can be substituted by a variety of techniques, e.g., site-directed mutagenesis, available in the art. Accordingly, any and all such variations of a sequence selected from the above table are a feature of the invention.

In addition to silent variations, other conservative variations that alter one, or a few amino acids in the encoded polypeptide, can be made without altering the function of the polypeptide, these conservative variants are, likewise, a feature of the invention.

For example, substitutions, deletions and insertions introduced into the sequences provided in the Sequence Listing are also envisioned by the invention. Such sequence modifications can be engineered into a sequence by site-directed mutagenesis (Wu (ed.) Meth. Enzymol. (1993) vol. 217, Academic Press) or the other methods noted below. Amino acid substitutions are typically of single residues; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. In preferred embodiments, deletions or insertions are made in adjacent pairs, e.g., a deletion of two residues or insertion of two residues. Substitutions, deletions, insertions or any combination thereof can be combined to arrive at a sequence. The mutations that are made in the polynucleotide encoding the transcription factor should not place the sequence out of reading frame and should not create complementary regions that could produce secondary mRNA structure. Preferably, the polypeptide encoded by the DNA performs the desired function.

Conservative substitutions are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the Table 2 when it is desired to maintain the activity of the protein. Table 2 shows amino acids which can be substituted for an amino acid in a protein and which are typically regarded as conservative substitutions.



**Table 2**

<b>Residue</b>	<b>Conservative Substitutions</b>
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Gln	Asn
Cys	Ser
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu, Val
Leu	Ile; Val
Lys	Arg; Gln
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr; Gly
Thr	Ser; Val
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Similar substitutions are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the Table 3 when it is desired to maintain the activity of the protein. Table 3 shows amino acids which can be substituted for an amino acid in a protein and which are typically regarded as structural and functional substitutions. For example, a residue in column 1 of Table 3 may be substituted with residue in column 2; in addition, a residue in column 2 of Table 3 may be substituted with the residue of column 1.

**Table 3**

<b>Residue</b>	<b>Similar Substitutions</b>
Ala	Ser; Thr; Gly; Val; Leu; Ile
Arg	Lys; His; Gly
Asn	Gln; His; Gly; Ser; Thr
Asp	Glu, Ser; Thr
Gln	Asn; Ala
Cys	Ser; Gly
Glu	Asp
Gly	Pro; Arg
His	Asn; Gln; Tyr; Phe; Lys; Arg
Ile	Ala; Leu; Val; Gly; Met
Leu	Ala; Ile; Val; Gly; Met
Lys	Arg; His; Gln; Gly; Pro
Met	Leu; Ile; Phe
Phe	Met; Leu; Tyr; Trp; His; Val; Ala
Ser	Thr; Gly; Asp; Ala; Val; Ile; His
Thr	Ser; Val; Ala; Gly
Trp	Tyr; Phe; His
Tyr	Trp; Phe; His
Val	Ala; Ile; Leu; Gly; Thr; Ser; Glu

Substitutions that are less conservative than those in Table 2 can be selected by picking residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in protein properties will be those in which (a) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a

hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine.

### **VIII. Further Modifying Sequences of the Invention – Mutation/Forced Evolution**

In addition to generating silent or conservative substitutions as noted, above, the present invention optionally includes methods of modifying the sequences of the Sequence Listing. In the methods, nucleic acid or protein modification methods are used to alter the given sequences to produce new sequences and/or to chemically or enzymatically modify given sequences to change the properties of the nucleic acids or proteins.

Thus, in one embodiment, given nucleic acid sequences are modified, e.g., according to standard mutagenesis or artificial evolution methods to produce modified sequences. The modified sequences may be created using purified natural polynucleotides isolated from any organism or may be synthesized from purified compositions and chemicals using chemical means well known to those of skill in the art. For example, Ausubel, *supra*, provides additional details on mutagenesis methods. Artificial forced evolution methods are described, for example, by Stemmer (1994) Nature 370:389-391, Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751, and U.S. Patents 5,811,238, 5,837,500, and 6,242,568. Methods for engineering synthetic transcription factors and other polypeptides are described, for example, by Zhang et al. (2000) J. Biol. Chem. 275:33850-33860, Liu et al. (2001) J. Biol. Chem. 276:11323-11334, and Isalan et al. (2001) Nature Biotechnol. 19:656-660. Many other mutation and evolution methods are also available and expected to be within the skill of the practitioner.

Similarly, chemical or enzymatic alteration of expressed nucleic acids and polypeptides can be performed by standard methods. For example, sequence can be modified by addition of lipids, sugars, peptides, organic or inorganic compounds, by

the inclusion of modified nucleotides or amino acids, or the like. For example, protein modification techniques are illustrated in Ausubel, *supra*. Further details on chemical and enzymatic modifications can be found herein. These modification methods can be used to modify any given sequence, or to modify any sequence produced by the various mutation and artificial evolution modification methods noted herein.

Accordingly, the invention provides for modification of any given nucleic acid by mutation, evolution, chemical or enzymatic modification, or other available methods, as well as for the products produced by practicing such methods, e.g., using the sequences herein as a starting substrate for the various modification approaches.

For example, optimized coding sequence containing codons preferred by a particular prokaryotic or eukaryotic host can be used e.g., to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced using a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, preferred stop codons for *Saccharomyces cerevisiae* and mammals are TAA and TGA, respectively. The preferred stop codon for monocotyledonous plants is TGA, whereas insects and *E. coli* prefer to use TAA as the stop codon.

The polynucleotide sequences of the present invention can also be engineered in order to alter a coding sequence for a variety of reasons, including but not limited to, alterations which modify the sequence to facilitate cloning, processing and/or expression of the gene product. For example, alterations are optionally introduced using techniques which are well known in the art, e.g., site-directed mutagenesis, to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to introduce splice sites, etc.

Furthermore, a fragment or domain derived from any of the polypeptides of the invention can be combined with domains derived from other transcription factors or synthetic domains to modify the biological activity of a transcription factor. For instance, a DNA-binding domain derived from a transcription factor of the invention

can be combined with the activation domain of another transcription factor or with a synthetic activation domain. A transcription activation domain assists in initiating transcription from a DNA-binding site. Examples include the transcription activation region of VP16 or GAL4 (Moore et al. (1998) Proc. Natl. Acad. Sci. USA 95: 376-381; and Aoyama et al. (1995) Plant Cell 7:1773-1785), peptides derived from bacterial sequences (Ma and Ptashne (1987) Cell 51; 113-119) and synthetic peptides (Giniger and Ptashne, (1987) Nature 330:670-672).

#### **IX. Expression and Modification of Polypeptides**

Typically, polynucleotide sequences of the invention are incorporated into recombinant DNA (or RNA) molecules that direct expression of polypeptides of the invention in appropriate host cells, transgenic plants, in vitro translation systems, or the like. Due to the inherent degeneracy of the genetic code, nucleic acid sequences which encode substantially the same or a functionally equivalent amino acid sequence can be substituted for any listed sequence to provide for cloning and expressing the relevant homologue.

#### **X. Vectors, Promoters, and Expression Systems**

The present invention includes recombinant constructs comprising one or more of the nucleic acid sequences herein. The constructs typically comprise a vector, such as a plasmid, a cosmid, a phage, a virus (e.g., a plant virus), a bacterial artificial chromosome (BAC), a yeast artificial chromosome (YAC), or the like, into which a nucleic acid sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available.

General texts that describe molecular biological techniques useful herein, including the use and production of vectors, promoters and many other relevant topics, include Berger, Sambrook and Ausubel, *supra*. Any of the identified sequences can be incorporated into a cassette or vector, e.g., for expression in plants. A number of expression vectors suitable for stable transformation of plant cells or for the establishment of transgenic plants have been described including those described in

Weissbach and Weissbach, (1989) Methods for Plant Molecular Biology, Academic Press, and Gelvin et al., (1990) Plant Molecular Biology Manual, Kluwer Academic Publishers. Specific examples include those derived from a Ti plasmid of *Agrobacterium tumefaciens*, as well as those disclosed by Herrera-Estrella et al. (1983) Nature 303: 209, Bevan (1984) Nucl Acid Res. 12: 8711-8721, Klee (1985) Bio/Technology 3: 637-642, for dicotyledonous plants.

Alternatively, non-Ti vectors can be used to transfer the DNA into monocotyledonous plants and cells by using free DNA delivery techniques. Such methods can involve, for example, the use of liposomes, electroporation, microprojectile bombardment, silicon carbide whiskers, and viruses. By using these methods transgenic plants such as wheat, rice (Christou (1991) Bio/Technology 9: 957-962) and corn (Gordon-Kamm (1990) Plant Cell 2: 603-618) can be produced. An immature embryo can also be a good target tissue for monocots for direct DNA delivery techniques by using the particle gun (Weeks et al. (1993) Plant Physiol 102: 1077-1084; Vasil (1993) Bio/Technology 10: 667-674; Wan and Lemeaux (1994) Plant Physiol 104: 37-48, and for *Agrobacterium*-mediated DNA transfer (Ishida et al. (1996) Nature Biotech 14: 745-750).

Typically, plant transformation vectors include one or more cloned plant coding sequence (genomic or cDNA) under the transcriptional control of 5' and 3' regulatory sequences and a dominant selectable marker. Such plant transformation vectors typically also contain a promoter (e.g., a regulatory region controlling inducible or constitutive, environmentally-or developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, an RNA processing signal (such as intron splice sites), a transcription termination site, and/or a polyadenylation signal.

Examples of constitutive plant promoters which can be useful for expressing the TF sequence include: the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant tissues (*see, e.g.,* Odell et al. (1985) Nature 313:810-812); the nopaline synthase promoter (An et al. (1988) Plant Physiol 88:547-552); and the octopine synthase promoter (Fromm et al. (1989) Plant Cell 1: 977-984).

A variety of plant gene promoters that regulate gene expression in response to environmental, hormonal, chemical, developmental signals, and in a tissue-active manner can be used for expression of a TF sequence in plants. Choice of a promoter is based largely on the phenotype of interest and is determined by such factors as tissue (e.g., seed, fruit, root, pollen, vascular tissue, flower, carpel, etc.), inducibility (e.g., in response to wounding, heat, cold, drought, light, pathogens, etc.), timing, developmental stage, and the like. Numerous known promoters have been characterized and can favorably be employed to promote expression of a polynucleotide of the invention in a transgenic plant or cell of interest. For example, tissue specific promoters include: seed-specific promoters (such as the napin, phaseolin or DC3 promoter described in US Pat. No. 5,773,697), fruit-specific promoters that are active during fruit ripening (such as the *dru 1* promoter (US Pat. No. 5,783,393), or the 2A11 promoter (US Pat. No. 4,943,674) and the tomato polygalacturonase promoter (Bird et al. (1988) Plant Mol Biol 11:651), root-specific promoters, such as those disclosed in US Patent Nos. 5,618,988, 5,837,848 and 5,905,186, pollen-active promoters such as PTA29, PTA26 and PTA13 (US Pat. No. 5,792,929), promoters active in vascular tissue (Ringli and Keller (1998) Plant Mol Biol 37:977-988), flower-specific (Kaiser et al. (1995) Plant Mol Biol 28:231-243), pollen (Baerson et al. (1994) Plant Mol Biol 26:1947-1959), carpels (Ohl et al. (1990) Plant Cell 2:837-848), pollen and ovules (Baerson et al. (1993) Plant Mol Biol 22:255-267), auxin-inducible promoters (such as that described in van der Kop et al. (1999) Plant Mol Biol 39:979-990 or Baumann et al. (1999) Plant Cell 11:323-334), cytokinin-inducible promoter (Guevara-Garcia (1998) Plant Mol Biol 38:743-753), promoters responsive to gibberellin (Shi et al. (1998) Plant Mol Biol 38:1053-1060, Willmott et al. (1998) 38:817-825) and the like. Additional promoters are those that elicit expression in response to heat (Ainley et al. (1993) Plant Mol Biol 22: 13-23), light (e.g., the pea *rbcS*-3A promoter, Kuhlemeier et al. (1989) Plant Cell 1:471, and the maize *rbcS* promoter, Schaffner and Sheen (1991) Plant Cell 3: 997); wounding (e.g., *wun1*, Siebertz et al. (1989) Plant Cell 1: 961); pathogens (such as the PR-1 promoter described in Buchel et al. (1999) Plant Mol. Biol. 40:387-396, and the PDF1.2 promoter described in Manners et al. (1998) Plant Mol. Biol. 38:1071-80), and chemicals such as methyl jasmonate or salicylic acid (Gatz et al. (1997) Plant Mol Biol 48: 89-108). In addition, the timing of the expression can be controlled by using

promoters such as those acting at senescence (An and Amazon (1995) Science 270: 1986-1988); or late seed development (Odell et al. (1994) Plant Physiol 106:447-458).

Plant expression vectors can also include RNA processing signals that can be positioned within, upstream or downstream of the coding sequence. In addition, the expression vectors can include additional regulatory sequences from the 3'-untranslated region of plant genes, e.g., a 3' terminator region to increase mRNA stability of the mRNA, such as the PI-II terminator region of potato or the octopine or nopaline synthase 3' terminator regions.

### **Additional Expression Elements**

Specific initiation signals can aid in efficient translation of coding sequences. These signals can include, e.g., the ATG initiation codon and adjacent sequences. In cases where a coding sequence, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence (e.g., a mature protein coding sequence), or a portion thereof, is inserted, exogenous transcriptional control signals including the ATG initiation codon can be separately provided. The initiation codon is provided in the correct reading frame to facilitate transcription. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers appropriate to the cell system in use.

### **Expression Hosts**

The present invention also relates to host cells which are transduced with vectors of the invention, and the production of polypeptides of the invention (including fragments thereof) by recombinant techniques. Host cells are genetically engineered (i.e., nucleic acids are introduced, e.g., transduced, transformed or transfected) with the vectors of this invention, which may be, for example, a cloning vector or an expression vector comprising the relevant nucleic acids herein. The vector is optionally a plasmid, a viral particle, a phage, a naked nucleic acid, *etc.* The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the relevant gene. The culture conditions, such as temperature, pH and the like, are those



previously used with the host cell selected for expression, and will be apparent to those skilled in the art and in the references cited herein, including, Sambrook and Ausubel.

The host cell can be a eukaryotic cell, such as a yeast cell, or a plant cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Plant protoplasts are also suitable for some applications. For example, the DNA fragments are introduced into plant tissues, cultured plant cells or plant protoplasts by standard methods including electroporation (Fromm et al., (1985) Proc. Natl. Acad. Sci. USA 82, 5824, infection by viral vectors such as cauliflower mosaic virus (CaMV) (Hohn et al., (1982) Molecular Biology of Plant Tumors, (Academic Press, New York) pp. 549-560; US 4,407,956), high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface (Klein et al., (1987) Nature 327, 70-73), use of pollen as vector (WO 85/01856), or use of *Agrobacterium tumefaciens* or *A. rhizogenes* carrying a T-DNA plasmid in which DNA fragments are cloned. The T-DNA plasmid is transmitted to plant cells upon infection by *Agrobacterium tumefaciens*, and a portion is stably integrated into the plant genome (Horsch et al. (1984) Science 233:496-498; Fraley et al. (1983) Proc. Natl. Acad. Sci. USA 80, 4803).

The cell can include a nucleic acid of the invention which encodes a polypeptide, wherein the cell expresses a polypeptide of the invention. The cell can also include vector sequences, or the like. Furthermore, cells and transgenic plants that include any polypeptide or nucleic acid above or throughout this specification, e.g., produced by transduction of a vector of the invention, are an additional feature of the invention.

For long-term, high-yield production of recombinant proteins, stable expression can be used. Host cells transformed with a nucleotide sequence encoding a polypeptide of the invention are optionally cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The protein or fragment thereof produced by a recombinant cell may be secreted, membrane-bound, or contained intracellularly, depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing

polynucleotides encoding mature proteins of the invention can be designed with signal sequences which direct secretion of the mature polypeptides through a prokaryotic or eukaryotic cell membrane.

#### **XI. Modified Amino Acid Residues**

Polypeptides of the invention may contain one or more modified amino acid residues. The presence of modified amino acids may be advantageous in, for example, increasing polypeptide half-life, reducing polypeptide antigenicity or toxicity, increasing polypeptide storage stability, or the like. Amino acid residue(s) are modified, for example, co-translationally or post-translationally during recombinant production or modified by synthetic or chemical means.

Non-limiting examples of a modified amino acid residue include incorporation or other use of acetylated amino acids, glycosylated amino acids, sulfated amino acids, prenylated (e.g., farnesylated, geranylgeranylated) amino acids, PEG modified (e.g., "PEGylated") amino acids, biotinylated amino acids, carboxylated amino acids, phosphorylated amino acids, etc. References adequate to guide one of skill in the modification of amino acid residues are replete throughout the literature.

The modified amino acid residues may prevent or increase affinity of the polypeptide for another molecule, including, but not limited to, polynucleotide, proteins, carbohydrates, lipids and lipid derivatives, and other organic or synthetic compounds.

#### **XII. Identification of Additional Factors**

A transcription factor provided by the present invention can also be used to identify additional endogenous or exogenous molecules that can affect a phenotype or trait of interest. On the one hand, such molecules include organic (small or large molecules) and/or inorganic compounds that affect expression of (i.e., regulate) a particular transcription factor. Alternatively, such molecules include endogenous molecules that are acted upon either at a transcriptional level by a transcription factor of the invention to modify a phenotype as desired. For example, the transcription factors can be employed to identify one or more downstream gene with which is subject to a regulatory effect of the transcription factor. In one approach, a

transcription factor or transcription factor homologue of the invention is expressed in a host cell, e.g., a transgenic plant cell, tissue or explant, and expression products, either RNA or protein, of likely or random targets are monitored, e.g., by hybridization to a microarray of nucleic acid probes corresponding to genes expressed in a tissue or cell type of interest, by two-dimensional gel electrophoresis of protein products, or by any other method known in the art for assessing expression of gene products at the level of RNA or protein. Alternatively, a transcription factor of the invention can be used to identify promoter sequences (i.e., binding sites) involved in the regulation of a downstream target. After identifying a promoter sequence, interactions between the transcription factor and the promoter sequence can be modified by changing specific nucleotides in the promoter sequence or specific amino acids in the transcription factor that interact with the promoter sequence to alter a plant trait. Typically, transcription factor DNA-binding sites are identified by gel shift assays. After identifying the promoter regions, the promoter region sequences can be employed in double-stranded DNA arrays to identify molecules that affect the interactions of the transcription factors with their promoters (Bulyk et al. (1999) Nature Biotechnology 17:573-577).

The identified transcription factors are also useful to identify proteins that modify the activity of the transcription factor. Such modification can occur by covalent modification, such as by phosphorylation, or by protein-protein (homo or heteropolymer) interactions. Any method suitable for detecting protein-protein interactions can be employed. Among the methods that can be employed are co-immunoprecipitation, cross-linking and co-purification through gradients or chromatographic columns, and the two-hybrid yeast system.

The two-hybrid system detects protein interactions in vivo and is described in Chien et al. ((1991), Proc. Natl. Acad. Sci. USA 88:9578-9582) and is commercially available from Clontech (Palo Alto, Calif.). In such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to the TF polypeptide and the other consists of the transcription activator protein's activation domain fused to an unknown protein that is encoded by a cDNA that has been recombined into the plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are

transformed into a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (e.g., lacZ) whose regulatory region contains the transcription activator's binding site. Either hybrid protein alone cannot activate transcription of the reporter gene. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product. Then, the library plasmids responsible for reporter gene expression are isolated and sequenced to identify the proteins encoded by the library plasmids. After identifying proteins that interact with the transcription factors, assays for compounds that interfere with the TF protein-protein interactions can be preformed.

### **XIII. Identification of Modulators**

In addition to the intracellular molecules described above, extracellular molecules that alter activity or expression of a transcription factor, either directly or indirectly, can be identified. For example, the methods can entail first placing a candidate molecule in contact with a plant or plant cell. The molecule can be introduced by topical administration, such as spraying or soaking of a plant, and then the molecule's effect on the expression or activity of the TF polypeptide or the expression of the polynucleotide monitored. Changes in the expression of the TF polypeptide can be monitored by use of polyclonal or monoclonal antibodies, gel electrophoresis or the like. Changes in the expression of the corresponding polynucleotide sequence can be detected by use of microarrays, Northern, quantitative PCR, or any other technique for monitoring changes in mRNA expression. These techniques are exemplified in Ausubel et al. (eds) Current Protocols in Molecular Biology, John Wiley & Sons (1998, and supplements through 2001). Such changes in the expression levels can be correlated with modified plant traits and thus identified molecules can be useful for soaking or spraying on fruit, vegetable and grain crops to modify traits in plants.

Essentially any available composition can be tested for modulatory activity of expression or activity of any nucleic acid or polypeptide herein. Thus, available libraries of compounds such as chemicals, polypeptides, nucleic acids and the like can be tested for modulatory activity. Often, potential modulator compounds can be dissolved in aqueous or organic (e.g., DMSO-based) solutions for easy delivery to the

cell or plant of interest in which the activity of the modulator is to be tested. Optionally, the assays are designed to screen large modulator composition libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays).

In one embodiment, high throughput screening methods involve providing a combinatorial library containing a large number of potential compounds (potential modulator compounds). Such "combinatorial chemical libraries" are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as target compounds.

A combinatorial chemical library can be, e.g., a collection of diverse chemical compounds generated by chemical synthesis or biological synthesis. For example, a combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (e.g., in one example, amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound of a set length). Exemplary libraries include peptide libraries, nucleic acid libraries, antibody libraries (see, e.g., Vaughn et al. (1996) Nature Biotechnology, 14(3):309-314 and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang et al. Science (1996) 274:1520-1522 and U.S. Patent 5,593,853), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), and small organic molecule libraries (see, e.g., benzodiazepines, Baum C&EN Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337) and the like.

Preparation and screening of combinatorial or other libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent 5,010,175; Furka, (1991) Int. J. Pept. Prot. Res. 37:487-493; and Houghton et al. (1991) Nature 354:84-88). Other chemistries for generating chemical diversity libraries can also be used.

In addition, as noted, compound screening equipment for high-throughput screening is generally available, e.g., using any of a number of well known robotic systems that have also been developed for solution phase chemistries useful in assay systems. These systems include automated workstations including an automated synthesis apparatus and robotic systems utilizing robotic arms. Any of the above devices are suitable for use with the present invention, e.g., for high-throughput screening of potential modulators. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art.

Indeed, entire high throughput screening systems are commercially available. These systems typically automate entire procedures including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. Similarly, microfluidic implementations of screening are also commercially available.

The manufacturers of such systems provide detailed protocols the various high throughput. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like. The integrated systems herein, in addition to providing for sequence alignment and, optionally, synthesis of relevant nucleic acids, can include such screening apparatus to identify modulators that have an effect on one or more polynucleotides or polypeptides according to the present invention.

In some assays it is desirable to have positive controls to ensure that the components of the assays are working properly. At least two types of positive controls are appropriate. That is, known transcriptional activators or inhibitors can be incubated with cells/plants/ etc. in one sample of the assay, and the resulting increase/decrease in transcription can be detected by measuring the resulting increase in RNA/ protein expression, etc., according to the methods herein. It will be appreciated that modulators can also be combined with transcriptional activators or inhibitors to find modulators that inhibit transcriptional activation or transcriptional repression. Either expression of the nucleic acids and proteins herein or any

additional nucleic acids or proteins activated by the nucleic acids or proteins herein, or both, can be monitored.

In an embodiment, the invention provides a method for identifying compositions that modulate the activity or expression of a polynucleotide or polypeptide of the invention. For example, a test compound, whether a small or large molecule, is placed in contact with a cell, plant (or plant tissue or explant), or composition comprising the polynucleotide or polypeptide of interest and a resulting effect on the cell, plant, (or tissue or explant) or composition is evaluated by monitoring, either directly or indirectly, one or more of: expression level of the polynucleotide or polypeptide, activity (or modulation of the activity) of the polynucleotide or polypeptide. In some cases, an alteration in a plant phenotype can be detected following contact of a plant (or plant cell, or tissue or explant) with the putative modulator, e.g., by modulation of expression or activity of a polynucleotide or polypeptide of the invention. Modulation of expression or activity of a polynucleotide or polypeptide of the invention may also be caused by molecular elements in a signal transduction second messenger pathway and such modulation can affect similar elements in the same or another signal transduction second messenger pathway.

#### XIV. Subsequences

Also contemplated are uses of polynucleotides, also referred to herein as oligonucleotides, typically having at least 12 bases, preferably at least 15, more preferably at least 20, 30, or 50 bases, which hybridize under at least highly stringent (or ultra-high stringent or ultra-ultra-high stringent conditions) conditions to a polynucleotide sequence described above. The polynucleotides may be used as probes, primers, sense and antisense agents, and the like, according to methods as noted *supra*.

Subsequences of the polynucleotides of the invention, including polynucleotide fragments and oligonucleotides are useful as nucleic acid probes and primers. An oligonucleotide suitable for use as a probe or primer is at least about 15 nucleotides in length, more often at least about 18 nucleotides, often at least about 21 nucleotides, frequently at least about 30 nucleotides, or about 40 nucleotides, or more in length. A

nucleic acid probe is useful in hybridization protocols, e.g., to identify additional polypeptide homologues of the invention, including protocols for microarray experiments. Primers can be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR) or other nucleic-acid amplification methods. See Sambrook and Ausubel, *supra*.

In addition, the invention includes an isolated or recombinant polypeptide including a subsequence of at least about 15 contiguous amino acids encoded by the recombinant or isolated polynucleotides of the invention. For example, such polypeptides, or domains or fragments thereof, can be used as immunogens, e.g., to produce antibodies specific for the polypeptide sequence, or as probes for detecting a sequence of interest. A subsequence can range in size from about 15 amino acids in length up to and including the full length of the polypeptide.

To be encompassed by the present invention, an expressed polypeptide which comprises such a polypeptide subsequence performs at least one biological function of the intact polypeptide in substantially the same manner, or to a similar extent, as does the intact polypeptide. For example, a polypeptide fragment can comprise a recognizable structural motif or functional domain such as a DNA binding domain that binds to a specific DNA promoter region, an activation domain or a domain for protein-protein interactions.

## **XV. Production of Transgenic Plants**

### **Modification of Traits**

The polynucleotides of the invention are favorably employed to produce transgenic plants with various traits, or characteristics, that have been modified in a desirable manner, e.g., to improve the seed characteristics of a plant. For example, alteration of expression levels or patterns (e.g., spatial or temporal expression patterns) of one or more of the transcription factors (or transcription factor homologues) of the invention, as compared with the levels of the same protein found in a wild type plant, can be used to modify a plant's traits. An illustrative example of



trait modification, improved characteristics, by altering expression levels of a particular transcription factor is described further in the Examples and the Sequence Listing.

#### Arabidopsis as a model system

*Arabidopsis thaliana* is the object of rapidly growing attention as a model for genetics and metabolism in plants. *Arabidopsis* has a small genome, and well documented studies are available. It is easy to grow in large numbers and mutants defining important genetically controlled mechanisms are either available, or can readily be obtained. Various methods to introduce and express isolated homologous genes are available (see Koncz, et al., eds. *Methods in Arabidopsis Research*. et al. (1992), World Scientific, New Jersey, New Jersey, in "Preface"). Because of its small size, short life cycle, obligate autogamy and high fertility, *Arabidopsis* is also a choice organism for the isolation of mutants and studies in morphogenetic and development pathways, and control of these pathways by transcription factors (Koncz, *supra*, p. 72). A number of studies introducing transcription factors into *A. thaliana* have demonstrated the utility of this plant for understanding the mechanisms of gene regulation and trait alteration in plants. See, for example, Koncz, *supra*, and U.S. Patent Number 6,417,428).

#### Arabidopsis genes in transgenic plants.

Expression of genes which encode transcription factors modify expression of endogenous genes, polynucleotides, and proteins are well known in the art. In addition, transgenic plants comprising isolated polynucleotides encoding transcription factors may also modify expression of endogenous genes, polynucleotides, and proteins. Examples include Peng et al. (1997, *Genes and Development* 11:3194-3205) and Peng et al. (1999, *Nature*, 400:256-261). In addition, many others have demonstrated that an *Arabidopsis* transcription factor expressed in an exogenous plant species elicits the same or very similar phenotypic response. See, for example, Fu et al. (2001, *Plant Cell* 13:1791-1802); Nandi et al. (2000, *Curr. Biol.* 10:215-218); Coupland (1995, *Nature* 377:482-483); and Weigel and Nilsson (1995, *Nature* 377:482-500).

Homologous genes introduced into transgenic plants.

Homologous genes that may be derived from any plant, or from any source whether natural, synthetic, semi-synthetic or recombinant, and that share significant sequence identity or similarity to those provided by the present invention, may be introduced into plants, for example, crop plants, to confer desirable or improved traits. Consequently, transgenic plants may be produced that comprise a recombinant expression vector or cassette with a promoter operably linked to one or more sequences homologous to presently disclosed sequences. The promoter may be, for example, a plant or viral promoter.

The invention thus provides for methods for preparing transgenic plants, and for modifying plant traits. These methods include introducing into a plant a recombinant expression vector or cassette comprising a functional promoter operably linked to one or more sequences homologous to presently disclosed sequences. Plants and kits for producing these plants that result from the application of these methods are also encompassed by the present invention.

The complete descriptions of the traits associated with each polynucleotide of the invention is fully disclosed in Table 4, Table 5, and Table 6.

Table 4

Polynucleotide SEQ ID NO.	GID No.	Trait	Category	Family	Comment	Polypeptide SEQ ID NO.	Conserved Domains
1	G1322	Cold	Abiotic stress	MYB-(R1)R2R3	Increased seedling vigor	2	(26-130)
3	G2130	Cold; heat; Botrytis	Abiotic stress; disease	AP2	Increased susceptibility to chilling; better germination in heat; Increased susceptibility to Botrytis	4	(93-160)
5	G256	Cold	Abiotic stress	MYB-(R1)R2R3	Better germination and growth in cold	6	(13-115)
7	G394	Cold	Abiotic stress	HB	More sensitive to chilling.	8	(121-182)
9	G664	Cold	Abiotic stress	MYB-(R1)R2R3	Better germination and growth in cold	10	(13-116)
11	G864	Cold; heat	Abiotic stress	AP2	Adult stage chilling sensitivity; better germination in heat	12	(119-186)
13	G1820	Drought; osmotic; hormone sensitivity	Abiotic stress; hormone sensitivity	CAAT	Increased tolerance to drought; better germination in high salt; reduced ABA sensitivity	14	(70-133)
15	G912	Drought; freezing	Abiotic stress	AP2	Increased survival in drought conditions, freezing tolerant	16	(51-118)
17	G2583	Drought; low humidity; insect resistance	Abiotic stress; disease	AP2	Tolerance to drought and/or low humidity; resistance to insects	18	(4-71)
19	G975	Drought; low humidity; insect resistance	Abiotic stress; disease	AP2	Tolerance to drought and/or low humidity; resistance to insects	20	(4-71)
21	G1322	Drought; low humidity; insect resistance	Abiotic stress; disease	MYB-(R1)R2R3	Tolerance to drought and/or low humidity; resistance to insects	22	(26-130)
23	G303	Drought; low humidity; insect resistance	Abiotic stress; disease	HLH/MYC	Tolerance to drought and/or low humidity; resistance to insects	24	(92-161)
25	G720	Freezing	Abiotic stress	GARP	Knockout: increased susceptibility to freezing (knockout); overexpressor: slightly more freezing tolerant	26	(301-349)
27	G913	Freezing	Abiotic stress	AP2	Increased tolerance to freezing	28	(62-128)
29	G1305	Heat	Abiotic stress	MYB-(R1)R2R3	Reduced chlorosis	30	(15-118)
31	G1645	Heat	Abiotic stress	MYB-(R1)R2R3	Reduced germination vigor	32	(90-210)
33	G1841	Heat	Abiotic stress	AP2	Better germination under heat stress	34	(83-150)
35	G2430	Heat	Abiotic stress	GARP	Increased tolerance	36	(425-478)
37	G3	Heat	Abiotic stress	AP2	More sensitive	38	(28-95)
39	G464	Heat	Abiotic stress	IAA	Better germination and growth in heat	40	(20-28, 71-82, 126-142, 187-224)
41	G682	Heat	Abiotic stress	MYB-related	Better germination and growth in heat	42	(27-63)
43	G964	Heat	Abiotic stress	HB	Better germination and growth	44	(126-186)
45	G1792	Nutrient uptake; Botrytis; Erysiphe; Fusarium	Abiotic stress; disease	AP2	Increased tolerance to nitrogen-limited medium; increased fungal disease resistance	46	(17-85)
47	G1794	Nutrient uptake; osmotic	Abiotic stress	AP2	Reduced root growth; increased sensitivity to osmotic stress	48	(182-248)
49	G1946	Nutrient uptake	Abiotic stress	HS	Increased root growth on phosphate-free media	50	(32-130)
51	G225	Nutrient uptake	Abiotic stress	MYB-related	Increased tolerance to nitrogen-limited medium	52	(39-76)
53	G226	Nutrient uptake; sodium chloride	Abiotic stress	MYB-related	Increased tolerance to nitrogen-limited medium; susceptible to high salt	54	(28-78)
55	G419	Nutrient uptake	Abiotic stress	HB	Increased tolerance to potassium-free medium	56	(392-452)

Table 4

57	G545	Nutrient uptake; sodium chloride; Erysiphe; Fusarium; Pseudomonas	Abiotic stress; disease	Z-C2H2	Increased tolerance to phosphate-free medium, high salt; increased susceptibility to disease	58	(82-102, 136-154)
59	G561	Nutrient uptake	Abiotic stress	bZIP	Increased tolerance to potassium-free medium	60	(248-308)
61	G911	Nutrient uptake	Abiotic stress	RING/C3H2C3	Increased growth on potassium-free medium	62	(86-129)
63	G1069	Osmotic; hormone sensitivity	Abiotic stress; hormone sensitivity	AT-hook	Better germination under osmotic stress; reduced ABA sensitivity	64	(67-74)
65	G1089	Osmotic	Abiotic stress	BZIPT2	Better germination under osmotic stress	66	(425-500)
67	G1452	Osmotic; hormone sensitivity	Abiotic stress; hormone sensitivity	NAC	Better germination on sucrose and salt; reduced sensitivity to ABA	68	(30-177)
69	G175	Osmotic	Abiotic stress	WRKY	Increased tolerance	70	(178-234, 372-428)
71	G1852	Osmotic	Abiotic stress	AKR	Better root growth	72	(1-801)
73	G1863	Osmotic	Abiotic stress	GRF-like	Decreased germination under salt stress	74	(TBD)
75	G188	Osmotic; Fusarium	Abiotic stress; disease	WRKY	Better germination under osmotic stress; increased susceptibility	76	(175-222)
77	G1930	Osmotic	Abiotic stress	AP2	Better germination under osmotic stress	78	(59-124)
79	G2053	Osmotic	Abiotic stress	NAC	Increased root growth	80	(10-149)
81	G2140	Osmotic; hormone sensitivity	Abiotic stress; hormone sensitivity	HLH/MYC	Better germination on NaCl and sucrose; decreased sensitivity to ABA	82	(167-242)
83	G2153	Osmotic	Abiotic stress	AT-hook	Better germination	84	(75-94, 162-206)
85	G2379	Osmotic	Abiotic stress	TH	Increased seedling vigor on high sucrose media	86	(19-110, 173-232)
87	G2701	Osmotic	Abiotic stress	MYB-related	Better germination on high NaCl and sucrose	88	(33-81, 129-183)
89	G2719	Osmotic	Abiotic stress	MYB-(R1)R2R3	Increased seedling vigor on high sucrose	90	(56-154)
91	G2789	Osmotic; hormone sensitivity	Abiotic stress; hormone sensitivity	AT-hook	Better germination on high sucrose; reduced ABA sensitivity	92	(53-73, 121-165)
93	G303	Osmotic	Abiotic stress	HLH/MYC	Better germination on high sucrose and high NaCl	94	(92-161)
95	G325	Osmotic	Abiotic stress	Z-CO-like	Slightly better germination on high sucrose and high NaCl	96	(5-28, 48-71)
97	G353	Osmotic	Abiotic stress	Z-C2H2	Increased seedling vigor on PEG	98	(41-61, 84-104)
99	G47	Osmotic	Abiotic stress	AP2	Better root growth	100	(11-80)
101	G489	Osmotic	Abiotic stress	CAAT	Increased tolerance to osmotic stress	102	(57-156)
103	G502	Osmotic	Abiotic stress	NAC	Increased sensitivity to osmotic stress	104	(10-155)
105	G526	Osmotic	Abiotic stress	NAC	Increased sensitivity to osmotic stress	106	(21-149)
107	G921	Osmotic	Abiotic stress	WRKY	Increased sensitivity to osmotic stress	108	(146-203)
109	G922	Osmotic; sodium chloride	Abiotic stress	SCR	Better germination on high sucrose; better germination and increased root growth on high salt	110	(225-242)
111	G926	Osmotic; hormone sensitivity	Abiotic stress; hormone sensitivity	CAAT	Increased tolerance to osmotic stress (salt and sucrose); reduced sensitivity to ABA	112	(131-225)
113	G481	Osmotic	Abiotic stress	CAAT	Overexpressors more tolerant to high sucrose in a germination assay	114	(20-109)
115	G1807	Oxidative	Abiotic stress	bZIP	More sensitive to acifluorfen	116	(249-287)

Table 4

117	G477	Oxidative; Sclerotinia	Abiotic stress; disease	SBP	Increased sensitivity to oxidative stress; Increased susceptibility to Sclerotinia-caused disease	118	(108-233)
119	G789	Oxidative; Sclerotinia	Abiotic stress; disease	HLH/MYC	Increased sensitivity to oxidative stress; Increased susceptibility to Sclerotinia-caused disease	120	(253-313)
121	G1836	Sodium chloride	Abiotic stress	CAAT	Better germination in high salt	122	(30-164)
123	G196	Sodium chloride	Abiotic stress	WRKY	Increased tolerance	124	(223-283)
125	G2110	Sodium chloride	Abiotic stress	WRKY	Increased tolerance	126	(239-298)
127	G22	Sodium chloride	Abiotic stress	AP2	Increased tolerance to high salt	128	(89-157)
129	G2713	Sodium chloride	Abiotic stress	TUBBY	Increased root growth	130	(59-270, 400-445)
131	G312	Sodium chloride	Abiotic stress	SCR	Slightly better germination on high NaCl	132	(320-336)
133	G482	Sodium chloride	Abiotic stress	CAAT	Increased tolerance to high salt	134	(25-116)
135	G801	Sodium chloride	Abiotic stress	PCF	Better germination on NaCl	136	(32-93)
137	G867	Sodium chloride	Abiotic stress	AP2	Better seedling vigor	138	(59-124)
139	G884	Sodium chloride	Abiotic stress	WRKY	Increased root growth	140	(227-285, 407-465)
141	G2133	Glyphosate	Herbicide sensitivity	AP2	Increased tolerance	142	(11-83)
143	G2517	Glyphosate	Herbicide sensitivity	WRKY	Increased tolerance	144	(TBD)
145	G343	Glyphosate	Herbicide sensitivity	GATA/Zn	Increase in resistance	146	(178-214)
147	G1082	Hormone sensitivity	Hormone sensitivity	HLH/MYC	Altered response to the growth hormone ethylene	148	(308-359)
149	G1095	Hormone sensitivity	Hormone sensitivity	RING/C3H2C3	Increased sensitivity to ACC	150	(134-159)
151	G1134	Hormone sensitivity	Hormone sensitivity	HLH/MYC	Altered response to the growth hormone ethylene	152	(198-247)
153	G12	Hormone sensitivity	Hormone sensitivity	AP2	Increased sensitivity to ACC	154	(TBD)
155	G1330	Hormone sensitivity	Hormone sensitivity	MYB-(R1)R2R3	Altered response to ethylene	156	(28-134)
157	G1666	Hormone sensitivity	Hormone sensitivity	HLH/MYC	Increased sensitivity to ACC	158	(353-420)
159	G546	Hormone sensitivity	Hormone sensitivity	RING/C3H2C3	Decreased sensitivity to ABA	160	(114-155)
161	G760	Hormone sensitivity	Hormone sensitivity	NAC	Hypersensitive to ACC	162	(12-156)
163	G913	Hormone sensitivity	Hormone sensitivity	AP2	Increased sensitivity to ACC	164	(62-128)
165	G1064	Botrytis	Disease	PCF	Increased sensitivity	166	(116-179)
167	G1084	Botrytis	Disease	BZIP2	Increased susceptibility	168	(TBD)
169	G1196	Botrytis	Disease	AKR	Increased susceptibility	170	(179-254)
171	G1255	Botrytis	Disease	Z-CO-like	Increased susceptibility	172	(18-56)
173	G1514	Botrytis	Disease	RING/C3HC4	Increased susceptibility	174	(20-50)
175	G1756	Botrytis	Disease	WRKY	Increased susceptibility	176	(TBD)
177	G1880	Botrytis	Disease	Z-C2H2	Increased resistance	178	(69-89, 111-139)
179	G1919	Botrytis	Disease	RING/C3HC4	Increased tolerance	180	(214-287)
181	G1936	Botrytis; Sclerotinia	Disease	PCF	Increased tolerance	182	(64-129)
183	G1950	Botrytis	Disease	AKR	Increased tolerance	184	(65-228)
185	G2069	Botrytis	Disease	bZIP	Increased susceptibility	186	(TBD)
187	G207	Botrytis	Disease	MYB-(R1)R2R3	Increased susceptibility	188	(6-106)
189	G2380	Botrytis	Disease	TH	Increased susceptibility	190	(107-181)
191	G248	Botrytis	Disease	MYB-(R1)R2R3	Increased susceptibility	192	(264-332)
193	G2555	Botrytis	Disease	HLH/MYC	Increased susceptibility	194	(175-245)

Table 4

195	G28	Botrytis; Sclerotinia; Erysiphe	Disease	AP2	Botrytis, Sclerotinia; Increased tolerance; Erysiphe:	196	(145-213)
197	G371	Botrytis	Disease	RING/C3HC4	Increased resistance	198	(21-74)
199	G812	Botrytis	Disease	HS	Increased susceptibility	200	(31-120)
201	G865	Botrytis	Disease	AP2	Increased tolerance	202	(36-103)
203	G940	Botrytis	Disease	EIL	Increased susceptibility	204	(86-98)
205	G558	Defense gene expression	Disease	bZIP	Increased expression of defense genes	206	(45-105)
207	G569	Defense gene expression	Disease	bZIP	Decreased expression of defense genes	208	(90-153)
209	G1266	Erysiphe	Disease	AP2	Increased tolerance	210	(79-147)
211	G19	Erysiphe	Disease	AP2	Increased tolerance; repressed by methyl jasmonate and induced by ACC	212	(76-145)
213	G237	Erysiphe	Disease	MYB-(R1)R2R3	Increased tolerance	214	(11-113)
215	G378	Erysiphe	Disease	RING/C3H2C3	Increased resistance	216	(196-237)
217	G409	Erysiphe	Disease	HB	Increased tolerance	218	(64-124)
219	G591	Erysiphe	Disease	HLH/MYC	Increased tolerance	220	(143-240)
221	G616	Erysiphe	Disease	TEO	Increased tolerance	222	(39-95)
223	G869	Erysiphe	Disease	AP2	Increased tolerance	224	(109-177)
225	G881	Erysiphe	Disease	WRKY	Increased susceptibility	226	(176-233)
227	G1047	Fusarium	Disease	bZIP	Increased tolerance	228	(129-180)
229	G1363	Fusarium	Disease	CAAT	Increased tolerance	230	(174-226)
231	G147	Fusarium	Disease	MADS	Increased susceptibility	232	(2-57)
233	G896	Fusarium	Disease	Z-LSLlike	Increased susceptibility	234	(18-39)
235	G418	Pseudomonas	Disease	HB	Increased tolerance	236	(500-560)
237	G525	Pseudomonas	Disease	NAC	Increased tolerance	238	(23-167)
239	G1927	Sclerotinia	Disease	NAC	Increased tolerance	240	(17-188)
241	G278	Sclerotinia	Disease	AKR	Increased susceptibility	242	(2-593)
243	G594	Sclerotinia	Disease	HLH/MYC	Increased susceptibility	244	(140-204)
245	G805	Sclerotinia	Disease	PCF	Increased susceptibility	246	(51-114)

Table 5

SEQ ID NO	GID	Test Sequence ID	Smallest Sum. Probability	Test Sequence Species	Test Sequence GenBank Annotation
25	G720	BG450227	3.20E-31	[Medicago truncatula]	NF015E11DT1F1087 Drought Medicago trunc
25	G720	AF318580	3.60E-26	[Zea mays]	putative transcription factor ZmGLK1 (G1k1) mRNA,
25	G720	AF318581	1.50E-22	[Oryza sativa]	putative transcription factor OsGLK1 (G1k1) mR
25	G720	AW906970	5.50E-20	[Solanum tuberosum]	EST343197 potato stolon, Cornell Universi
25	G720	BE330069	2.20E-15	[Glycine max]	so73a08.y1 Gm-c1040 Glycine max cDNA clone GENO
25	G720	AI896489	4.20E-15	[Lycopersicon esculentum]	EST265932 tomato callus, TAMU Lycop
25	G720	BI070070	4.70E-14	[Populus tremula x Populus tremuloides]	C013P55U Populus stra
25	G720	AW618051	3.10E-11	[Lycopersicon pennellii]	EST314101 L. pennellii trichome, Cor
25	G720	BG049455	1.80E-09	[Sorghum bicolor]	OV1_20 B11.g1_A002 Ovary 1 (OV1) Sorghum bl
25	G720	AW719680	1.20E-08	[Lotus japonicus]	LJNEST8C3r Lotus japonicus nodule library 5
25	G720	gi11177540	1.30E-28	[Zea mays]	putative transcription factor Golden2.
25	G720	gi13940498	6.90E-28	[Oryza sativa]	putative transcription factor OsGLK1.
25	G720	gi4519671	5.60E-05	[Nicotiana tabacum]	transfactor.
25	G720	gi6942190	0.0022	[Mesembryanthemum crystallinum]	CDPK substrate protein 1; C
25	G720	gi5916207	0.0053	[Chlamydomonas reinhardtii]	regulatory protein of P-starvat
25	G720	gi1247386	0.23	[Nicotiana glauca]	PRP2.
25	G720	gi169878	0.24	[Sesbania rostrata]	nodulin.
25	G720	gi1808688	0.3	[Sporobolus stapfianus]	hypothetical protein.
25	G720	gi99948	0.61	[Glycine max]	proline-rich protein 2 precursor - soybean.
25	G720	gi82091	0.7	[Lycopersicon esculentum]	hydroxyproline-rich glycoprotein
45	G1792	AI776626	1.20E-32	[Lycopersicon esculentum]	EST257726 tomato resistant, Cornell
45	G1792	BF646324	1.30E-29	[Medicago truncatula]	NF074E05EC1F1038 Elicited cell culture
45	G1792	BM178875	4.60E-29	[Glycine max]	sal60f01.y1 Gm-c1072 Glycine max cDNA clone SOY
45	G1792	AC025907	4.20E-27	[Oryza sativa]	chromosome 10 clone nbxb0094K20, *** SEQUENCIN
45	G1792	AP004623	1.90E-22	[Oryza sativa (japonica cultivar-group)]	( ) chromosome 8 clo
45	G1792	BM268956	2.10E-22	[Zea mays]	MEST402-H11.univ (SUM5-RN Zea mays cDNA clone MEST
45	G1792	AF245119	3.80E-22	[Mesembryanthemum crystallinum]	AP2-related transcription fac
45	G1792	BH517030	2.10E-21	[Brassica oleracea]	BOHRB76TF BOHR Brassica oleracea genomic
45	G1792	BM437083	5.00E-21	[Vitis vinifera]	VVA014A06_53661 An expressed sequence tag da
45	G1792	BF275652	5.90E-21	[Gossypium arboreum]	GA_Eb0024J23f Gossypium arboreum 7-10 d
45	G1792	gi1732406	1.00E-24	[Nicotiana tabacum]	S25-XP1 DNA binding protein.
45	G1792	gi12597874	1.70E-24	[Oryza sativa]	putative ethylene-responsive element binding
45	G1792	gi7528276	3.60E-24	[Mesembryanthemum crystallinum]	AP2-related transcription f
45	G1792	gi2213783	3.30E-23	[Lycopersicon esculentum]	Pti5.
45	G1792	gi19034045	8.60E-23	[Oryza sativa (japonica cultivar-group)]	putative DNA bindi
45	G1792	gi8980313	8.60E-23	[Catharanthus roseus]	AP2-domain DNA-binding protein.
45	G1792	gi8809571	8.60E-23	[Nicotiana sylvestris]	ethylene-responsive element binding

Table 5

45	G1792	gi17385636	5.50E-21	[Matricaria chamomilla]	ethylene-responsive element binding
45	G1792	gi8571476	5.00E-20	[Atriplex hortensis]	apetala2 domain-containing protein.
45	G1792	gi1688233	1.30E-19	[Solanum tuberosum]	DNA binding protein homolog.
175	G1756	AV423663	6.50E-34	[Lotus japonicus]	AV423663 Lotus japonicus young plants (two-
175	G1756	AW596933	3.10E-30	[Glycine max]	sj84f07.y1 Gm-c1034 Glycine max cDNA clone GENO
175	G1756	AW447931	8.40E-24	[Triticum aestivum]	BRY 1082 BRY Triticum aestivum cDNA clone
175	G1756	BI266533	1.50E-22	[Medicago truncatula]	NF100A09IN1F1097 Insect herbivory Medic
175	G1756	AW218278	1.10E-19	[Lycopersicon esculentum]	EST303459 tomato radicle, 5 d post-
175	G1756	AP002744	2.60E-18	[Oryza sativa]	genomic DNA, chromosome 1, PAC clone:P0006C01,
175	G1756	BG600477	3.00E-15	[Solanum tuberosum]	EST505372 cSTS Solanum tuberosum cDNA clo
175	G1756	gi11761072	7.10E-23	[Oryza sativa]	hypothetical protein.
175	G1756	gi4322940	3.10E-15	[Nicotiana tabacum]	DNA-binding protein 2.
175	G1756	gi1432056	9.00E-12	[Petroselinum crispum]	WRKY3.
175	G1756	gi4894963	2.00E-11	[Avena sativa]	DNA-binding protein WRKY3.
175	G1756	gi927025	7.40E-11	[Cucumis sativus]	SPF1-like DNA-binding protein.
175	G1756	gi13620227	2.50E-10	[Lycopersicon esculentum]	hypothetical protein.
175	G1756	gi11993901	6.10E-10	[Dactylis glomerata]	somatic embryogenesis related protein.
175	G1756	gi1159877	7.20E-10	[Avena fatua]	DNA-binding protein.
175	G1756	gi484261	8.60E-10	[Ipomoea batatas]	SPF1 protein.
175	G1756	gi3420906	1.30E-09	[Pimpinella brachycarpa]	zinc finger protein; WRKY1.
113	G481	BG440251	9.00E-42	[Gossypium arboreum]	GA_Ea0006K20f Gossypium arboreum 7-10 d
113	G481	BM887558	3.90E-41	[Glycine max]	sam40c09.y1 Gm-c1068 Glycine max cDNA clone SOY
113	G481	ZMNFYB	1.70E-40	[Zea mays]	Z.mays mRNA for CAAT-box DNA binding protein subun
113	G481	AI728916	2.40E-40	[Gossypium hirsutum]	BNLGH12022 Six-day Cotton fiber Gossyp
113	G481	AW775623	3.80E-40	[Medicago truncatula]	EST334688 DSIL Medicago truncatula cDNA
113	G481	AW738727	9.80E-40	[Lycopersicon esculentum]	EST340154 tomato flower buds, anthe
113	G481	BG599785	1.70E-38	[Solanum tuberosum]	EST504680 cSTS Solanum tuberosum cDNA clo
113	G481	BE413647	2.50E-38	[Triticum aestivum]	SCU001.E10.R990714.ITEC SCU Wheat Endospe
113	G481	BF065056	5.80E-38	[Hordeum vulgare]	HV_CEB0022M01f Hordeum vulgare seedling gre
113	G481	BG314203	1.00E-37	[Triticum monococcum]	WHE2460_E10_I20ZS Triticum monococcum l
113	G481	gi22380	1.10E-45	[Zea mays]	CAAT-box DNA binding protein subunit B (NF-YB).
113	G481	gi15408794	2.60E-30	[Oryza sativa]	putative CCAAT-binding transcription factor
113	G481	gi16902054	1.00E-28	[Vernonia galamensis]	CCAAT-box binding factor HAP3 B domai
113	G481	gi16902050	2.70E-28	[Glycine max]	CCAAT-box binding factor HAP3 B domain.
113	G481	gi16902056	4.30E-28	[Argemone mexicana]	CCAAT-box binding factor HAP3 B domain.
113	G481	gi16902058	1.50E-23	[Triticum aestivum]	CCAAT-box binding factor HAP3 B domain.
141	G2133	BE320193	1.20E-22	[Medicago truncatula]	NF024B04RT1F1029 Developing root Medica
141	G2133	AP003379	2.10E-20	[Oryza sativa]	chromosome 1 clone P0408G07. *** SEQUENCING IN
141	G2133	BG332521	1.50E-16	[Pinus taeda]	NXPV_073_H03_F_NXPV (Nsf Xylem Planings wood Ve



Table 5

141	G2133	BG046836	1.50E-15	[Glycine max]	saa62d08.y1 Gm-c1060 Glycine max cDNA clone GEN
141	G2133	BG321374	1.60E-15	[Descurainia sophia]	Ds01_06d08 R Ds01_AAFc ECORC cold stress
141	G2133	BE432190	4.00E-15	[Lycopersicon esculentum]	EST398719 tomato breaker fruit, TIG
141	G2133	BI431368	6.70E-15	[Solanum tuberosum]	EST534129 P. infestans-challenged leaf So
141	G2133	AU085418	1.20E-14	[Cryptomeria japonica]	AU085418 Cryptomeria japonica Inner ba
141	G2133	BG446456	2.80E-14	[Gossypium arboreum]	GA_Eb0034M18f Gossypium arboreum 7-10 d
141	G2133	AI728590	3.10E-14	[Gossypium hirsutum]	BNLGH11133 Six-day Cotton fiber Gossyp
141	G2133	gi8571476	2.50E-16	[Atriplex hortensis]	apetala2 domain-containing protein.
141	G2133	gi14140155	3.30E-15	[Oryza sativa]	putative AP2 domain transcription factor.
141	G2133	gi5616086	3.00E-14	[Brassica napus]	dehydration responsive element binding pro
141	G2133	gi12225916	1.10E-13	[Zea mays]	unnamed protein product.
141	G2133	gi10798644	4.20E-13	[Nicotiana tabacum]	AP2 domain-containing transcription fac
141	G2133	gi8980313	1.90E-12	[Catharanthus roseus]	AP2-domain DNA-binding protein.
141	G2133	gi2213785	1.30E-11	[Lycopersicon esculentum]	Pt6.
141	G2133	gi6478845	7.30E-11	[Matricaria chamomilla]	ethylene-responsive element binding
141	G2133	gi12658319	1.20E-10	[Hordeum vulgare]	CBF1-like protein BCBF1.
141	G2133	gi8809573	1.50E-10	[Nicotiana sylvestris]	ethylene-responsive element binding
143	G2517	BI470088	1.50E-35	[Glycine max]	saf38f09.y3 Gm-c1077 Glycine max cDNA clone GEN
143	G2517	BI208323	2.60E-31	[Lycopersicon esculentum]	EST526363 cTOS Lycopersicon esculen
143	G2517	BE445081	4.40E-28	[Triticum aestivum]	WHE1131_B06_D11ZS Wheat etiolated seedlin
143	G2517	AV408330	2.40E-26	[Lotus japonicus]	AV408330 Lotus japonicus young plants (two-
143	G2517	BE362650	7.10E-26	[Sorghum bicolor]	DG1_88_H02.b1_A002 Dark Grown 1 (DG1) Sorgh
143	G2517	BI308031	4.70E-25	[Medicago truncatula]	EST529441 GPOD Medicago truncatula cDNA
143	G2517	AP002839	2.00E-23	[Oryza sativa]	genomic DNA, chromosome 1, PAC clone:P0688A04,
143	G2517	BE216050	4.60E-21	[Hordeum vulgare]	HV_CEB0009E04f Hordeum vulgare seedling gre
143	G2517	BG889690	8.70E-21	[Solanum tuberosum]	EST515541 cSTD Solanum tuberosum cDNA clo
143	G2517	L35779	7.50E-19	[Brassica rapa]	BNAESTG Mustard flower buds Brassica rapa cDN
143	G2517	gi11761085	2.10E-25	[Oryza sativa]	putative DNA-binding protein homolog.
143	G2517	gi9187622	1.10E-20	[Solanum tuberosum]	WRKY DNA binding protein.
143	G2517	gi4322940	6.70E-17	[Nicotiana tabacum]	DNA-binding protein 2.
143	G2517	gi1432058	5.40E-15	[Petroselinum crispum]	WRKY2.
143	G2517	gi1159877	5.10E-14	[Avena fatua]	DNA-binding protein.
143	G2517	gi3420906	5.80E-13	[Pinipella brachycarpa]	zinc finger protein, WRKY1.
143	G2517	gi13620227	1.50E-12	[Lycopersicon esculentum]	hypothetical protein.
143	G2517	gi927025	6.40E-12	[Cucumis sativus]	SPF1-like DNA-binding protein.
143	G2517	gi4894963	1.40E-11	[Avena sativa]	DNA-binding protein WRKY3.
143	G2517	gi6723482	4.70E-11	[Betula pendula]	wrky-type DNA binding protein.
81	G2140	AI488313	4.80E-59	[Lycopersicon esculentum]	EST246635 tomato ovary, TAMU Lycopers
81	G2140	BE020519	6.10E-55	[Glycine max]	sm44g03.y1 Gm-c1028 Glycine max cDNA clone GENO

Table 5

81	G2140	AU093196	9.70E-46	[Oryza sativa subsp. japonica]	AU093196 Rice callus Oryza sat
81	G2140	AP003908	8.00E-39	[Oryza sativa]	chromosome 8 clone OJ1300 C09, *** SEQUENCING
81	G2140	BF647687	5.30E-37	[Medicago truncatula]	NF025A04EC1F1024 Elicited cell culture
81	G2140	AI054433	1.70E-26	[Mesembryanthemum crystallinum]	R6-R97 Ice plant Lambda Unl-Z
81	G2140	BE922838	1.70E-24	[Solanum tuberosum]	EST426607 potato leaves and petioles Sola
81	G2140	AU085144	1.80E-16	[Cryptomeria japonica]	AU085144 Cryptomeria japonica inner ba
81	G2140	BG275442	1.90E-13	[Pinus taeda]	NXSI_141 F07 F NXSI (Nsf Xylem Side wood Incln
81	G2140	BE051300	6.30E-13	[Zea mays]	za74h04.b50 Maize Glume cDNAs Library Zea mays cDN
81	G2140	gi8570062	1.70E-27	[Oryza sativa]	ESTs C26093(C11622),AU090634(C12429) corresp
81	G2140	gi527655	0.00014	[Pennisetum glaucum]	myc-like regulatory R gene product.
81	G2140	gi22472	0.0055	[Zea mays]	R-S protein (AA 1-612).
81	G2140	gi527665	0.017	[Sorghum bicolor]	myc-like regulatory R gene product.
81	G2140	gi1086526	0.058	[Oryza australiensis]	transcriptional activator Ra homolog.
81	G2140	gi1086534	0.081	[Oryza officinalis]	transcriptional activator Ra homolog.
81	G2140	gi10998404	0.15	[Petunia x hybrida]	anthocyanin 1.
81	G2140	gi527663	0.15	[Tripsacum australe]	myc-like regulatory R gene product.
81	G2140	gi1086536	0.4	[Oryza rufipogon]	transcriptional activator Ra homolog.
81	G2140	gi1086530	0.49	[Oryza longistaminata]	transcriptional activator Ra homolog
49	G1946	LPHSF8	1.10E-119	[Lycopersicon peruvianum]	L.peruvianum Lp-hsf8 mRNA for heat
49	G1946	AC087771	4.10E-112	[Medicago truncatula]	clone 8D15, *** SEQUENCING IN PROGRESS
49	G1946	LEHSF8	5.90E-103	[Lycopersicon esculentum]	L.esculentum Le-hsf8 gene for heat
49	G1946	AW569138	3.10E-75	[Glycine max]	sl63g09.y1 Gm-r1030 Glycine max cDNA clone GENO
49	G1946	BG890899	1.30E-70	[Solanum tuberosum]	EST516750 cSTD Solanum tuberosum cDNA clo
49	G1946	AC027658	4.60E-53	[Oryza sativa]	subsp. japonica BAC nbxb0006113, chromosome 10
49	G1946	AV833112	4.90E-52	[Hordeum vulgare subsp. vulgare]	AV833112 K. Sato unpublished
49	G1946	gi19492	2.80E-121	[Lycopersicon peruvianum]	heat shock transcription factor 8
49	G1946	gi19260	5.10E-106	[Lycopersicon esculentum]	heat stress transcription factor
49	G1946	gi662924	2.00E-47	[Glycine max]	heat shock transcription factor 21.
49	G1946	gi5821138	9.70E-46	[Nicotiana tabacum]	heat shock factor.
49	G1946	gi11761077	2.90E-40	[Oryza sativa]	putative heat shock factor protein 1 (HSF 1)
49	G1946	gi866742	3.20E-40	[Zea mays]	heat shock factor.
49	G1946	gi7158882	2.70E-38	[Medicago sativa]	heat shock transcription factor.
49	G1946	gi3550588	1.90E-30	[Pisum sativum]	heat shock transcription factor (HSFA).
49	G1946	gi100546	0.46	[Avena sativa]	avenin precursor - oat.
49	G1946	gi14190783	1	[Aplum graveolens]	putative phloem transcription factor M1.
71	G1852	AF220204	2.00E-95	[Malus domestica]	unknown mRNA.
71	G1852	BG597959	3.30E-95	[Solanum tuberosum]	EST496637 cSTS Solanum tuberosum cDNA clo
71	G1852	AC077693	4.50E-93	[Oryza sativa]	chromosome 10 clone OSJNBa0095C07, *** SEQUENC
71	G1852	BG445922	5.90E-91	[Gossypium arboreum]	GA_Ea0030A23f Gossypium arboreum 7-10 d

Table 5

71	G1852	BG581705	1.60E-90	[Medicago truncatula]	EST483440 GVN Medicago truncatula cDNA
71	G1852	BF009089	5.30E-86	[Glycine max]	ss73404.y1 Gm-c1062 Glycine max cDNA clone GENO
71	G1852	BE434670	2.20E-71	[Lycopersicon esculentum]	EST405748 tomato breaker fruit, TIG
71	G1852	BG317894	5.90E-56	[Pinus taeda]	NXPV_006_H10_F_NXPV (Nsf Xylem Planings wood Va
71	G1852	AV422410	1.20E-55	[Lotus japonicus]	AV422410 Lotus japonicus young plants (two-
71	G1852	AI055434	1.20E-50	[Gossypium hirsutum]	coau0003P22 Cotton Boll Abscission Zone
71	G1852	gi5734619	1.30E-103	[Oryza sativa]	Similar to Arabidopsis thaliana BAC F15P23 (
71	G1852	gi6752888	1.80E-98	[Malus x domestica]	unknown.
71	G1852	gi498042	0.81	[Senecio odoratus]	ORF.
71	G1852	gi4432741	1	[Dioscorea tenuipes]	phosphoglucose isomerase.
95	G325	AB001888	9.20E-43	[Oryza sativa]	mRNA for zinc finger protein, complete cds.
95	G325	BE558327	5.40E-30	[Hordeum vulgare]	HV_GEB0017D19f Hordeum vulgare seedling gre
95	G325	BG644908	1.80E-29	[Medicago truncatula]	EST508527 KV3 Medicago truncatula cDNA
95	G325	BG605313	3.10E-29	[Triticum aestivum]	WHE2331_C04_F07ZS Wheat pre-anthesis splk
95	G325	BG459023	1.60E-28	[Zea mays]	947052H08.y1 947 - 2 week shoot from Barkan lab Ze
95	G325	BG590826	3.20E-28	[Solanum tuberosum]	EST498668 P. Infestans-challenged leaf So
95	G325	BG412316	4.40E-27	[Sorghum bicolor]	OV2_40_D10.b1_A002 Ovary 2 (OV2) Sorghum bl
95	G325	BG127619	1.50E-25	[Lycopersicon esculentum]	EST473181 tomato shoot/meristem Lyc
95	G325	AW399335	7.70E-22	[Lycopersicon pennellii]	EST309835 L. pennellii trichome, Cor
95	G325	AW780799	3.90E-21	[Glycine max]	si76c08.y1 Gm-c1027 Glycine max cDNA clone GENO
95	G325	gi3618320	5.60E-49	[Oryza sativa]	zinc finger protein.
95	G325	gi3341723	1.40E-13	[Raphanus sativus]	CONSTANS-like 1 protein.
95	G325	gi11037311	5.00E-13	[Brassica nigra]	constans-like protein.
95	G325	gi2303683	6.00E-13	[Brassica napus]	unnamed protein product.
95	G325	gi4091806	2.80E-12	[Malus x domestica]	CONSTANS-like protein 2.
95	G325	gi10946337	7.00E-11	[Ipomoea nil]	CONSTANS-like protein.
95	G325	gi4557093	6.50E-10	[Pinus radiata]	zinc finger protein.
17	G2583	AW928465	1.40E-43	[Lycopersicon esculentum]	EST337253 tomato flower buds 8 mm t
17	G2583	BE023297	2.40E-42	[Glycine max]	sm80a10.y1 Gm-c1015 Glycine max cDNA clone GENO
17	G2583	AP003615	1.60E-30	[Oryza sativa]	chromosome 6 clone P0486H12, *** SEQUENCING IN
17	G2583	AU088998	2.90E-21	[Lotus japonicus]	AU088998 Lotus japonicus flower bud cDNA Lo
17	G2583	AT001828	4.60E-20	[Brassica rapa subsp. pekinensis]	AT001828 Flower bud cDNA Br
17	G2583	BG415973	2.40E-18	[Hordeum vulgare]	HVSMK0009E06f Hordeum vulgare testa/perica
17	G2583	BF647090	3.80E-17	[Medicago truncatula]	NF007A06EC1F1038 Elicited cell culture
17	G2583	BG560598	2.80E-16	[Sorghum propinquum]	RHIZ2_59_D07.b1_A003 Rhizome2 (RHIZ2) So
17	G2583	AW011200	6.60E-16	[Pinus taeda]	ST17H08 Pine TriPEX shoot tip library Pinus ta
17	G2583	BF479478	1.60E-15	[Mesembryanthemum crystallinum]	L48-3155T3 Ice plant Lambda U
17	G2583	gi19507	1.40E-16	[Lupinus polyphylus]	put. pPLZ2 product (AA 1-164).
17	G2583	gi10798644	1.00E-12	[Nicotiana tabacum]	AP2 domain-containing transcription fac

Table 5

17	G2583	gi8571476	4.70E-12	[Atriplex hortensis]	apetala2 domain-containing protein.
17	G2583	gi2213783	8.40E-12	[Lycopersicon esculentum]	Pti5.
17	G2583	gi809573	5.30E-11	[Nicotiana sylvestris]	ethylene-responsive element binding
17	G2583	gi4099914	8.40E-11	[Stylosanthes hamata]	ethylene-responsive element binding p
17	G2583	gi6478845	8.90E-11	[Matricaria chamomilla]	ethylene-responsive element binding
17	G2583	gi15290041	9.40E-11	[Oryza sativa]	hypothetical protein.
17	G2583	gi12225884	1.70E-10	[Zea mays]	unnamed protein product.
17	G2583	gi3264767	3.40E-10	[Prunus armeniaca]	AP2 domain containing protein.
21	G1322	AW032656	3.60E-43	[Lycopersicon esculentum]	EST276215 tomato callus, TAMU Lycop
21	G1322	BF649523	5.40E-42	[Medicago truncatula]	NF080G02EC1F-1019 Elicited cell culture
21	G1322	BG369720	1.50E-41	[Hordeum vulgare]	HVSME10025K10F Hordeum vulgare 20 DAP spike
21	G1322	BF325282	2.60E-41	[Glycine max]	su20e03.y1 Gm-c1066 Glycine max cDNA clone GENO
21	G1322	GPU33917	6.40E-41	[Cratogeomys plantagineum]	myb-related transcription factor
21	G1322	AW255388	2.40E-40	[Mentha x piperita]	ML407 peppermint glandular trichome Menth
21	G1322	AB052230	1.30E-39	[Arabis gemmifera]	gene for MYB transcription factor Atmyb2.
21	G1322	AY026332	9.40E-39	[Oryza sativa]	Myb transcription factor JAMYb mRNA, complete
21	G1322	PSMYB26	1.10E-37	[Pisum sativum]	P. sativum mRNA for Myb-like protein (Myb26).
21	G1322	AI055122	4.30E-37	[Gossypium hirsutum]	coau0003B15 Colton Boll Abscission Zone
21	G1322	gi1002796	5.50E-49	[Cratogeomys plantagineum]	Cpm10.
21	G1322	gi13177578	3.00E-46	[Oryza sativa]	Myb transcription factor JAMYb.
21	G1322	gi1841475	1.50E-39	[Pisum sativum]	Myb26.
21	G1322	gi14249015	9.00E-39	[Gossypium hirsutum]	myb-like transcription factor Myb 5.
21	G1322	gi82306	3.10E-37	[Antirrhinum majus]	myb protein 305 - garden snapdragon.
21	G1322	gi8247759	7.40E-34	[Triticum aestivum]	GAMYb protein.
21	G1322	gi2130046	2.30E-33	[Hordeum vulgare]	MybHV5 protein - barley.
21	G1322	gi15082210	7.50E-33	[Fragaria x ananassa]	transcription factor MYB1.
21	G1322	gi5139806	1.10E-31	[Glycine max]	GmMYB29A2.
21	G1322	gi1732247	2.90E-31	[Nicotiana tabacum]	transcription factor Myb1.
23	G303	BE021887	5.70E-28	[Glycine max]	sm63g05.y1 Gm-c1028 Glycine max cDNA clone GENO
23	G303	BI480474	8.60E-28	[Triticum aestivum]	WHE2903 F02 L03Z5 Wheat aluminum-stressed
23	G303	AC079935	5.40E-27	[Oryza sativa]	clone OSJNBa0095C06, *** SEQUENCING IN PROGRES
23	G303	AW573949	1.50E-22	[Medicago truncatula]	EST316540 GVN Medicago truncatula cDNA
23	G303	BF200153	1.30E-21	[Triticum monococcum]	WHE2252 F02 K04ZE Triticum monococcum s
23	G303	CAR011013	2.20E-18	[Cicer arietinum]	epicotyl EST, clone Can133.
23	G303	BE357342	2.60E-11	[Sorghum bicolor]	DG1_148 D06.g1 A002 Dark Grown 1 (DG1) Sorg
23	G303	BG275269	3.10E-11	[Pinus taeda]	NXSI_138 F10 F NXSI (Nsf Xylem Side wood Incln
23	G303	BF588006	7.50E-10	[Sorghum propinquum]	FM1_35 E08.g1 A003 Floral-Induced Merist
23	G303	AI491136	2.00E-09	[Lycopersicon esculentum]	EST241845 tomato shoot, Cornell Lyc
23	G303	gi3641870	1.40E-22	[Cicer arietinum]	hypothetical protein.

Table 5

23	G303	gi1086538	9.60E-05	[Oryza rufipogon]	transcriptional activator Rb homolog.
23	G303	gi1086534	0.00082	[Oryza officinalis]	transcriptional activator Ra homolog.
23	G303	gi527661	0.0015	[Phyllostachys acuta]	myc-like regulatory R gene product.
23	G303	gi527653	0.005	[Pennisetum glaucum]	myc-like regulatory R gene product.
23	G303	gi22195	0.0078	[Zea mays]	regulatory protein.
23	G303	gi13161337	0.014	[Oryza sativa]	putative transcription activator.
23	G303	gi10998404	0.014	[Petunia x hybrida]	anthocyanin 1.
23	G303	gi1086526	0.052	[Oryza australiensis]	transcriptional activator Ra homolog.
23	G303	gi527665	0.072	[Sorghum bicolor]	myc-like regulatory R gene product.
239	G1927	BG350410	3.30E-63	[Solanum tuberosum]	091B07 Mature tuber lambda ZAP Solanum tu
239	G1927	BF066070	8.80E-48	[Hordeum vulgare]	HV_CEB0014M06f Hordeum vulgare seedling gre
239	G1927	AW736414	1.80E-47	[Medicago truncatula]	EST332428 KV3 Medicago truncatula cDNA
239	G1927	BG159075	7.10E-47	[Sorghum propinquum]	RHIZ2_17_E07.b1_A003 Rhizome2 (RHIZ2) So
239	G1927	BI543350	7.70E-39	[Beta vulgaris]	S1A_F6 Sugar Beet peroxide germination cDNA I
239	G1927	BE131060	9.20E-38	[Mesembryanthemum crystallinum]	L48-1010T3 Ice plant Lambda U
239	G1927	BI140602	1.00E-37	[Sorghum bicolor]	IP1_51_D12.b1_A002 Immature pannicle 1 (IP1
239	G1927	AI779001	4.30E-37	[Lycopersicon esculentum]	EST259880 tomato susceptible, Come
239	G1927	BF484725	6.40E-36	[Triticum aestivum]	WHE2319_A01_B01ZS Wheat pre-anthesis splk
239	G1927	BE449422	7.30E-35	[Lycopersicon hirsutum]	EST356181 L. hirsutum trichome, Come
239	G1927	gi7716952	4.10E-40	[Medicago truncatula]	NAC1.
239	G1927	gi6175246	1.10E-35	[Lycopersicon esculentum]	jasmonic acid 2.
239	G1927	gi15148914	4.70E-35	[Phaseolus vulgaris]	NAC domain protein NAC2.
239	G1927	gi1279640	3.30E-34	[Petunia x hybrida]	NAM.
239	G1927	gi6730938	4.70E-34	[Oryza sativa]	OsNAC4 protein.
239	G1927	gi14485513	4.10E-31	[Solanum tuberosum]	putative NAC domain protein.
239	G1927	gi6732156	4.20E-31	[Triticum monococcum]	unnamed protein product.
239	G1927	gi4218535	1.30E-30	[Triticum sp.]	GRAB1 protein.
239	G1927	gi4986349	6.30E-25	[Nicotiana tabacum]	NAC-domain protein.
239	G1927	gi2982275	2.60E-08	[Picea mariana]	ATAF1-like protein.

Traits of interest

Examples of some of the traits that may be desirable in plants, and that may be provided by transforming the plants with the presently disclosed sequences, are listed in Table 6.

Table 6. Genes, traits and utilities that affect plant characteristics

<u>Trait Category</u>	<u>Traits</u>	<u>Transcription factor genes that impact traits</u>	<u>Utility</u> <u>Gene effect on:</u>
Resistance and tolerance	Salt stress resistance	G22; G196; G226; G303; G312; G325; G353; G482; G545; G801; G867; G884; G922; G926; G1452; G1794; G1820; G1836; G1843; G1863; G2053; G2110; G2140; G2153; G2379; G2701; G2713; G2719; G2789	Germination rate, survivability, yield; extended growth range
	Osmotic stress resistance	G47; G175; G188; G303; G325; G353; G489; G502; G526; G921; G922; G926; G1069; G1089; G1452; G1794; G1930; G2140; G2153; G2379; G2701; G2719; G2789;	Germination rate, survivability, yield
	Cold stress resistance; cold germination	G256; G394; G664; G864; G1322; G2130	Germination, growth, earlier planting
	Tolerance to freezing	G303; G325; G353; G720; G912; G913; G1794; G2053; G2140; G2153; G2379; G2701; G2719; G2789	Survivability, yield, appearance, extended range
	Heat stress resistance	G3; G464; G682; G864; G964; G1305; G1645; G2130 G2430	Germination, growth, later planting
	Drought, low humidity resistance	G303; G325; G353; G720; G912; G926; G1452; G1794; G1820; G1843; G2053; G2140; G2153; G2379; G2583; G2701; G2719; G2789	Survivability, yield, extended range
	Radiation resistance	G1052	Survivability, vigor, appearance
	Decreased herbicide sensitivity	G343; G2133; G2517	Resistant to increased herbicide use
	Increased herbicide sensitivity	G374; G877; G1519	Use as a herbicide target

	Oxidative stress	G477; G789; G1807; G2133; G2517	Improved yield, appearance, reduced senescence
	Light response	G183; G354; G375; G1062; G1322; G1331; G1488; G1494; G1521; G1786; G1794; G2144; G2555;	Germination, growth, development, flowering time
Development, morphology	Overall plant architecture	G24; G27; G31; G33; G47; G147; G156; G160; G182; G187; G195; G196; G211; G221; G237; G280; G342; G352; G357; G358; G360; G362; G364; G365; G367; G373; G377; G396; G431; G447; G479; G546; G546; G551; G578; G580; G596; G615; G617; G620; G625; G638; G658; G716; G725; G727; G730; G740; G770; G858; G865; G869; G872; G904; G910; G912; G920; G939; G963; G977; G979; G987; G988; G993; G1007; G1010; G1014; G1035; G1046; G1049; G1062; G1069; G1070; G1076; G1089; G1093; G1127; G1131; G1145; G1229; G1246; G1304; G1318; G1320; G1330; G1331; G1352; G1354; G1360; G1364; G1379; G1384; G1399; G1415; G1417; G1442; G1453; G1454; G1459; G1460; G1471; G1475; G1477; G1487; G1487; G1492; G1499; G1499; G1531; G1540; G1543; G1543; G1544; G1548; G1584; G1587; G1588; G1589; G1636; G1642; G1747; G1749; G1749; G1751; G1752; G1763; G1766; G1767; G1778; G1789; G1790; G1791; G1793; G1794; G1795; G1800; G1806; G1811; G1835; G1836; G1838; G1839; G1843; G1853; G1855; G1865; G1881; G1882; G1883; G1884; G1891; G1896; G1898; G1902; G1904; G1906; G1913; G1914; G1925; G1929; G1930; G1954; G1958; G1965; G1976;	Vascular tissues, lignin content; cell wall content; appearance

		G2057; G2107; G2133; G2134; G2151; G2154; G2157; G2181; G2290; G2299; G2340; G2340; G2346; G2373; G2376; G2424; G2465; G2505; G2509; G2512; G2513; G2519; G2520; G2533; G2534; G2573; G2589; G2687; G2720; G2787; G2789; G2893	
	Size: increased stature	G189; G1073; G1435; G2430	
	Size: reduced stature or dwarfism	G3; G5; G21; G23; G39; G165; G184; G194; G258; G280; G340; G343; G353; G354; G362; G363; G370; G385; G396; G439; G440; G447; G450; G550; G557; G599; G636; G652; G670; G671; G674; G729; G760; G804; G831; G864; G884; G898; G900; G912; G913; G922; G932; G937; G939; G960; G962; G977; G991; G1000; G1008; G1020; G1023; G1053; G1067; G1075; G1137; G1181; G1198; G1228; G1266; G1267; G1275; G1277; G1309; G1311; G1314; G1317; G1322; G1323; G1326; G1332; G1334; G1367; G1381; G1382; G1386; G1421; G1488; G1494; G1537; G1545; G1560; G1586; G1641; G1652; G1655; G1671; G1750; G1756; G1757; G1782; G1786; G1794; G1839; G1845; G1879; G1886; G1888; G1933; G1939; G1943; G1944; G2011; G2094; G2115; G2130; G2132; G2144; G2145; G2147; G2156; G2294; G2313; G2344; G2431; G2510; G2517; G2521; G2893; G2893	Ornamental; small stature provides wind resistance; creation of dwarf varieties
	Fruit size and number	G362	Biomass, yield, cotton boll fiber density
	Flower structure, inflorescence	G47; G259; G353; G354; G671; G732; G988; G1000; G1063; G1140; G1326; G1449; G1543; G1560; G1587; G1645; G1947; G2108; G2143; G2893	Ornamental horticulture; production of saffron or other edible flowers
	Number and development of trichomes	G225; G226; G247; G362; G585; G634; G676; G682; G1014; G1332; G1452; G1795;	Resistance to pests and desiccation; essential oil



		G2105	production
	Seed size, color, and number	G156; G450; G584; G652; G668; G858; G979; G1040; G1062; G1145; G1255; G1494; G1531; G1534; G1594; G2105; G2114;	Yield
	Root development, modifications	G9; G1482; G1534; G1794; G1852; G2053; G2136; G2140	
	Modifications to root hairs	G225; G226	Nutrient, water uptake, pathogen resistance
	Apical dominance	G559; G732; G1255; G1275; G1411; G1488; G1635; G2452; G2509	Ornamental horticulture
	Branching patterns	G568; G988; G1548	Ornamental horticulture, knot reduction, improved windscreen
	Leaf shape, color, modifications	G375; G377; G428; G438; G447; G464; G557; G577; G599; G635; G671; G674; G736; G804; G903; G977; G921; G922; G1038; G1063; G1067; G1073; G1075; G1146; G1152; G1198; G1267; G1269; G1452; G1484; G1586; G1594; G1767; G1786; G1792; G1886; G2059; G2094; G2105; G2113; G2117; G2143; G2144; G2431; G2452; G2465; G2587; G2583; G2724;	Appealing shape or shiny leaves for ornamental agriculture, increased biomass or photosynthesis
	Silique	G1134	Ornamental
	Stem morphology	G47; G438; G671; G748; G988; G1000	Ornamental; digestibility
	Shoot modifications	G390; G391	Ornamental stem bifurcations
Disease, Pathogen Resistance	Bacterial	G211; G347; G367; G418; G525; G545; G578; G1049	Yield, appearance, survivability, extended range
	Fungal	G19; G28; G28; G28; G147; G188; G207; G211; G237; G248; G278; G347; G367; G371; G378; G409; G477; G545; G545; G558; G569; G578; G591; G594; G616; G789; G805; G812; G865; G869; G872; G881; G896;	Yield, appearance, survivability, extended range

		G940; G1047; G1049; G1064; G1084; G1196; G1255; G1266; G1363; G1514; G1756; G1792; G1792; G1792; G1792; G1880; G1919; G1919; G1927; G1927; G1936; G1936; G1950; G2069; G2130; G2380; G2380; G2555	
Nutrients	Increased tolerance to nitrogen-limited soils	G225; G226; G1792	
	Increased tolerance to phosphate-limited soils	G419; G545; G561; G1946	
	Increased tolerance to potassium-limited soils	G561; G911	
Hormonal	Hormone sensitivity	G12; G546; G926; G760; G913; G926; G1062; G1069; G1095; G1134; G1330; G1452; G1666; G1820; G2140; G2789	Seed dormancy, drought tolerance; plant form, fruit ripening
Seed biochemistry	Production of seed prenyl lipids, including tocopherol	G214; G259; G490; G652; G748; G883; G1052; G1328; G1930; G2509; G2520	Antioxidant activity, vitamin E
	Production of seed sterols	G20	Precursors for human steroid hormones; cholesterol modulators
	Production of seed glucosinolates	G353; G484; G674; G1272; G1506; G1897; G1946; G2113; G2117; G2155; G2290; G2340	Defense against insects; putative anticancer activity; undesirable in animal feeds
	Modified seed oil content	G162; G162; G180; G192; G241; G265; G286; G291; G427; G509; G519; G561; G567; G590; G818; G849; G892; G961; G974; G1063; G1143; G1190; G1198; G1226; G1229; G1323; G1451; G1471; G1478; G1496; G1526; G1543; G1640; G1644; G1646; G1672; G1677; G1750; G1765; G1777; G1793; G1838; G1902; G1946; G1948; G2059; G2123; G2138; G2139; G2343; G2792; G2830	Vegetable oil production; increased caloric value for animal feeds; lutein content
	Modified seed oil	G217; G504; G622; G778;	Heat stability,

	composition	G791; G861; G869; G938; G965; G1417; G2192	digestibility of seed oils
	Modified seed protein content	G162; G226; G241; G371; G427; G509; G567; G597; G732; G849; G865; G892; G963; G988; G1323; G1323; G1419; G1478; G1488; G1634; G1637; G1641; G1644; G1652; G1677; G1777; G1777; G1818; G1820; G1903; G1909; G1946; G1946; G1958; G2059; G2117; G2417; G2509	Reduced caloric value for humans
Leaf biochemistry	Production of flavonoids	G1666*	Ornamental pigment production; pathogen resistance; health benefits
	Production of leaf glucosinolates	G264; G353; G484; G652; G674; G681; G1069; G1198; G1322; G1421; G1657; G1794; G1897; G1946; G2115; G2117; G2144; G2155; G2155; G2340; G2512; G2520; G2552	Defense against insects; putative anticancer activity; undesirable in animal feeds
	Production of diterpenes	G229	Induction of enzymes involved in alkaloid biosynthesis
	Production of anthocyanin	G546	Ornamental pigment
	Production of leaf phytosterols, inc. stigmastanol, campesterol	G561; G2131; G2424	Precursors for human steroid hormones; cholesterol modulators
	Leaf fatty acid composition	G214; G377; G861; G962; G975; G987; G1266; G1337; G1399; G1465; G1512; G2136; G2147; G2192	Nutritional value; increase in waxes for disease resistance
	Production of leaf prenyl lipids, including tocopherol	G214; G259; G280; G652; G987; G1543; G2509; G2520	Antioxidant activity, vitamin E
Biochemistry, general	Production of miscellaneous secondary metabolites	G229; G663	
	Sugar, starch, hemicellulose composition,	G158; G211; G211; G237; G242; G274; G598; G1012; G1266; G1309; G1309; G1641;	Food digestibility, hemicellulose & pectin content;

		G1765; G1865; G2094; G2094; G2589; G2589	fiber content; plant tensile strength, wood quality, pathogen resistance, pulp production; tuber starch content
Sugar sensing	Plant response to sugars	G26; G38; G43; G207; G218; G241; G254; G263; G308; G536; G567; G567; G680; G867; G912; G956; G996; G1068; G1225; G1314; G1314; G1337; G1759; G1804; G2153; G2379	Photosynthetic rate, carbohydrate accumulation, biomass production, source-sink relationships, senescence
Growth, Reproduction	Plant growth rate and development	G447; G617; G674; G730; G917; G937; G1035; G1046; G1131; G1425; G1452; G1459; G1492; G1589; G1652; G1879; G1943; G2430; G2431; G2465; G2521	Faster growth, increased biomass or yield, improved appearance; delay in bolting
	Embryo development	G167	
	Seed germination rate	G979; G1792; G2130	Yield
	Plant, seedling vigor	G561; G2346	Survivability, yield
	Senescence; cell death	G571; G636; G878; G1050; G1463; G1749; G1944; G2130; G2155; G2340; G2383	Yield, appearance; response to pathogens;
	Modified fertility	G39; G340; G439; G470; G559; G615; G652; G671; G779; G962; G977; G988; G1000; G1063; G1067; G1075; G1266; G1311; G1321; G1326; G1367; G1386; G1421; G1453; G1471; G1453; G1560; G1594; G1635; G1750; G1947; G2011; G2094; G2113; G2115; G2130; G2143; G2147; G2294; G2510; G2893	Prevents or minimizes escape of the pollen of GMOs
	Early flowering	G147; G157; G180; G183; G183; G184; G185; G208; G227; G294; G390; G390; G390; G391; G391; G427; G427; G490; G565; G590; G592; G720; G789; G865; G898; G898; G989; G989; G1037; G1037; G1142; G1225; G1225; G1226; G1242; G1305;	Faster generation time; synchrony of flowering; potential for introducing new traits to single variety

		G1305; G1380; G1380; G1480; G1480; G1488; G1494; G1545; G1545; G1649; G1706; G1760; G1767; G1767; G1820; G1841; G1841; G1842; G1843; G1843; G1946; G1946; G2010; G2030; G2030; G2144; G2144; G2295; G2295; G2347; G2348; G2348; G2373; G2373; G2509; G2509; G2555; G2555	
	Delayed flowering	G8; G47; G192; G214; G234; G361; G362; G562; G568; G571; G591; G680; G736; G748; G859; G878; G910; G912; G913; G971; G994; G1051; G1052; G1073; G1079; G1335; G1435; G1452; G1478; G1789; G1804; G1865; G1865; G1895; G1900; G2007; G2133; G2155; G2291; G2465	Delayed time to pollen production of GMO plants; synchrony of flowering; increased yield
	Extended flowering phase	G1947	
	Flower and leaf development	G259; G353; G377; G580; G638 G652; G858; G869; G917; G922; G932; G1063; G1075; G1140; G1425; G1452; G1499; G1548; G1645; G1865; G1897; G1933; G2094; G2124; G2140; G2143; G2535; G2557	Ornamental applications; decreased fertility
	Flower abscission	G1897	Ornamental: longer retention of flowers

\* When co-expressed with G669 and G663

#### Significance of modified plant traits

Currently, the existence of a series of maturity groups for different latitudes represents a major barrier to the introduction of new valuable traits. Any trait (e.g. disease resistance) has to be bred into each of the different maturity groups separately, a laborious and costly exercise. The availability of single strain, which could be grown at any latitude, would therefore greatly increase the potential for introducing new traits to crop species such as soybean and cotton.

For many of the traits, listed in Table 6 and below, that may be conferred to plants, a single transcription factor gene may be used to increase or decrease, advance

or delay, or improve or prove deleterious to a given trait. For example, overexpression of a transcription factor gene that naturally occurs in a plant may cause early flowering relative to non-transformed or wild-type plants. By knocking out the gene, or suppressing the gene (with, for example, antisense suppression) the plant may experience delayed flowering. Similarly, overexpressing or suppressing one or more genes can impart significant differences in production of plant products, such as different fatty acid ratios. Thus, suppressing a gene that causes a plant to be more sensitive to cold may improve a plant's tolerance of cold.

Salt stress resistance. Soil salinity is one of the more important variables that determines where a plant may thrive. Salinity is especially important for the successful cultivation of crop plants, particular in many parts of the world that have naturally high soil salt concentrations, or where the soil has been over-utilized. Thus, presently disclosed transcription factor genes that provide increased salt tolerance during germination, the seedling stage, and throughout a plant's life cycle would find particular value for imparting survivability and yield in areas where a particular crop would not normally prosper.

Osmotic stress resistance. Presently disclosed transcription factor genes that confer resistance to osmotic stress may increase germination rate under adverse conditions, which could impact survivability and yield of seeds and plants.

Cold stress resistance. The potential utility of presently disclosed transcription factor genes that increase tolerance to cold is to confer better germination and growth in cold conditions. The germination of many crops is very sensitive to cold temperatures. Genes that would allow germination and seedling vigor in the cold would have highly significant utility in allowing seeds to be planted earlier in the season with a high rate of survivability. Transcription factor genes that confer better survivability in cooler climates allow a grower to move up planting time in the spring and extend the growing season further into autumn for higher crop yields.

Tolerance to freezing. The presently disclosed transcription factor genes that impart tolerance to freezing conditions are useful for enhancing the survivability and appearance of plants conditions or conditions that would otherwise cause extensive

cellular damage. Thus, germination of seeds and survival may take place at temperatures significantly below that of the mean temperature required for germination of seeds and survival of non-transformed plants. As with salt tolerance, this has the added benefit of increasing the potential range of a crop plant into regions in which it would otherwise succumb. Cold tolerant transformed plants may also be planted earlier in the spring or later in autumn, with greater success than with non-transformed plants.

Heat stress tolerance. The germination of many crops is also sensitive to high temperatures. Presently disclosed transcription factor genes that provide increased heat tolerance are generally useful in producing plants that germinate and grow in hot conditions, may find particular use for crops that are planted late in the season, or extend the range of a plant by allowing growth in relatively hot climates.

Drought, low humidity tolerance. Strategies that allow plants to survive in low water conditions may include, for example, reduced surface area or surface oil or wax production. A number of presently disclosed transcription factor genes increase a plant's tolerance to low water conditions and provide the benefits of improved survivability, increased yield and an extended geographic and temporal planting range.

Radiation resistance. Presently disclosed transcription factor genes have been shown to increase lutein production. Lutein, like other xanthophylls such as zeaxanthin and violaxanthin, are important in the protection of plants against the damaging effects of excessive light. Lutein contributes, directly or indirectly, to the rapid rise of non-photochemical quenching in plants exposed to high light. Increased tolerance of field plants to visible and ultraviolet light impacts survivability and vigor, particularly for recent transplants. Also affected are the yield and appearance of harvested plants or plant parts. Crop plants engineered with presently disclosed transcription factor genes that cause the plant to produce higher levels of lutein therefore would have improved photoprotection, leading to less oxidative damage and increase vigor, survivability and higher yields under high light and ultraviolet light conditions.

Decreased herbicide sensitivity. Presently disclosed transcription factor genes that confer resistance or tolerance to herbicides (e.g., glyphosate) may find use in providing means to increase herbicide applications without detriment to desirable plants. This would allow for the increased use of a particular herbicide in a local environment, with the effect of increased detriment to undesirable species and less harm to transgenic, desirable cultivars.

Increased herbicide sensitivity. Knockouts of a number of the presently disclosed transcription factor genes have been shown to be lethal to developing embryos. Thus, these genes are potentially useful as herbicide targets.

Oxidative stress. In plants, as in all living things, abiotic and biotic stresses induce the formation of oxygen radicals, including superoxide and peroxide radicals. This has the effect of accelerating senescence, particularly in leaves, with the resulting loss of yield and adverse effect on appearance. Generally, plants that have the highest level of defense mechanisms, such as, for example, polyunsaturated moieties of membrane lipids, are most likely to thrive under conditions that introduce oxidative stress (e.g., high light, ozone, water deficit, particularly in combination). Introduction of the presently disclosed transcription factor genes that increase the level of oxidative stress defense mechanisms would provide beneficial effects on the yield and appearance of plants. One specific oxidizing agent, ozone, has been shown to cause significant foliar injury, which impacts yield and appearance of crop and ornamental plants. In addition to reduced foliar injury that would be found in ozone resistant plant created by transforming plants with some of the presently disclosed transcription factor genes, the latter have also been shown to have increased chlorophyll fluorescence (Yu-Sen Chang et al. Bot. Bull. Acad. Sin. (2001) 42: 265-272).

Heavy metal tolerance. Heavy metals such as lead, mercury, arsenic, chromium and others may have a significant adverse impact on plant respiration. Plants that have been transformed with presently disclosed transcription factor genes that confer improved resistance to heavy metals, through, for example, sequestering or reduced uptake of the metals will show improved vigor and yield in soils with relatively high concentrations of these elements. Conversely, transgenic transcription factors may



also be introduced into plants to confer an increase in heavy metal uptake, which may benefit efforts to clean up contaminated soils.

Light response. Presently disclosed transcription factor genes that modify a plant's response to light may be useful for modifying a plant's growth or development, for example, photomorphogenesis in poor light, or accelerating flowering time in response to various light intensities, quality or duration to which a non-transformed plant would not similarly respond. Examples of such responses that have been demonstrated include leaf number and arrangement, and early flower bud appearances.

Overall plant architecture. Several presently disclosed transcription factor genes have been introduced into plants to alter numerous aspects of the plant's morphology. For example, it has been demonstrated that a number of transcription factors may be used to manipulate branching, such as the means to modify lateral branching, a possible application in the forestry industry. Transgenic plants have also been produced that have altered cell wall content, lignin production, flower organ number, or overall shape of the plants. Presently disclosed transcription factor genes transformed into plants may be used to affect plant morphology by increasing or decreasing internode distance, both of which may be advantageous under different circumstances. For example, for fast growth of woody plants to provide more biomass, or fewer knots, increased internode distances are generally desirable. For improved wind screening of shrubs or trees, or harvesting characteristics of, for example, members of the Gramineae family, decreased internode distance may be advantageous. These modifications would also prove useful in the ornamental horticulture industry for the creation of unique phenotypic characteristics of ornamental plants.

Increased stature. For some ornamental plants, the ability to provide larger varieties may be highly desirable. For many plants, including fruit-bearing trees or trees and shrubs that serve as view or wind screens, increased stature provides obvious benefits. Crop species may also produce higher yields on larger cultivars.

Reduced stature or dwarfism. Presently disclosed transcription factor genes that decrease plant stature can be used to produce plants that are more resistant to damage by wind and rain, or more resistant to heat or low humidity or water deficit. Dwarf plants are also of significant interest to the ornamental horticulture industry, and particularly for home garden applications for which space availability may be limited.

Fruit size and number. Introduction of presently disclosed transcription factor genes that affect fruit size will have desirable impacts on fruit size and number, which may comprise increases in yield for fruit crops, or reduced fruit yield, such as when vegetative growth is preferred (e.g., with bushy ornamentals, or where fruit is undesirable, as with ornamental olive trees).

Flower structure, inflorescence, and development. Presently disclosed transgenic transcription factors have been used to create plants with larger flowers or arrangements of flowers that are distinct from wild-type or non-transformed cultivars. This would likely have the most value for the ornamental horticulture industry, where larger flowers or interesting presentations generally are preferred and command the highest prices. Flower structure may have advantageous effects on fertility, and could be used, for example, to decrease fertility by the absence, reduction or screening of reproductive components. One interesting application for manipulation of flower structure, for example, by introduced transcription factors could be in the increased production of edible flowers or flower parts, including saffron, which is derived from the stigmas of *Crocus sativus*.

Number and development of trichomes. Several presently disclosed transcription factor genes have been used to modify trichome number and amount of trichome products in plants. Trichome glands on the surface of many higher plants produce and secrete exudates that give protection from the elements and pests such as insects, microbes and herbivores. These exudates may physically immobilize insects and spores, may be insecticidal or ant-microbial or they may act as allergens or irritants to protect against herbivores. Trichomes have also been suggested to decrease transpiration by decreasing leaf surface air flow, and by exuding chemicals that protect the leaf from the sun.

Seed size, color and number. The introduction of presently disclosed transcription factor genes into plants that alter the size or number of seeds may have a significant impact on yield, both when the product is the seed itself, or when biomass of the vegetative portion of the plant is increased by reducing seed production. In the case of fruit products, it is often advantageous to modify a plant to have reduced size or number of seeds relative to non-transformed plants to provide seedless or varieties with reduced numbers or smaller seeds. Presently disclosed transcription factor genes have also been shown to affect seed size, including the development of larger seeds. . Seed size, in addition to seed coat integrity, thickness and permeability, seed water content and by a number of other components including antioxidants and oligosaccharides, may affect seed longevity in storage. This would be an important utility when the seed of a plant is the harvested crops, as with, for example, peas, beans, nuts, etc. Presently disclosed transcription factor genes have also been used to modify seed color, which could provide added appeal to a seed product.

Root development, modifications. By modifying the structure or development of roots by transforming into a plant one or more of the presently disclosed transcription factor genes, plants may be produced that have the capacity to thrive in otherwise unproductive soils. For example, grape roots that extend further into rocky soils, or that remain viable in waterlogged soils, would increase the effective planting range of the crop. It may be advantageous to manipulate a plant to produce short roots, as when a soil in which the plant will be growing is occasionally flooded, or when pathogenic fungi or disease-causing nematodes are prevalent.

Modifications to root hairs. Presently disclosed transcription factor genes that increase root hair length or number potentially could be used to increase root growth or vigor, which might in turn allow better plant growth under adverse conditions such as limited nutrient or water availability.

Apical dominance. The modified expression of presently disclosed transcription factors that control apical dominance could be used in ornamental horticulture, for example, to modifying plant architecture.

Branching patterns. Several presently disclosed transcription factor genes have been used to manipulate branching, which could provide benefits in the forestry industry. For example, reduction in the formation of lateral branches could reduce knot formation. Conversely, increasing the number of lateral branches could provide utility when a plant is used as a windscreen, or may also provide ornamental advantages.

Leaf shape, color and modifications. It has been demonstrated in laboratory experiments that overexpression of some of the presently disclosed transcription factors produced marked effects on leaf development. At early stages of growth, these transgenic seedlings developed narrow, upward pointing leaves with long petioles, possibly indicating a disruption in circadian-clock controlled processes or nyctinastic movements. Other transcription factor genes can be used to increase plant biomass; large size would be useful in crops where the vegetative portion of the plant is the marketable portion.

Siliques. Genes that later silique conformation in brassicates may be used to modify fruit ripening processes in brassicates and other plants, which may positively affect seed or fruit quality.

Stem morphology and shoot modifications. Laboratory studies have demonstrated that introducing several of the presently disclosed transcription factor genes into plants can cause stem bifurcations in shoots, in which the shoot meristems split to form two or three separate shoots. This unique appearance would be desirable in ornamental applications.

Diseases, pathogens and pests. A number of the presently disclosed transcription factor genes have been shown to or are likely to confer resistance to various plant diseases, pathogens and pests. The offending organisms include fungal pathogens *Fusarium oxysporum*, *Botrytis cinerea*, *Sclerotinia sclerotiorum*, and *Erysiphe orontii*. Bacterial pathogens to which resistance may be conferred include *Pseudomonas syringae*. Other problem organisms may potentially include nematodes, mollicutes, parasites, or herbivorous arthropods. In each case, one or more transformed transcription factor genes may provide some benefit to the plant to

help prevent or overcome infestation. The mechanisms by which the transcription factors work could include increasing surface waxes or oils, surface thickness, local senescence, or the activation of signal transduction pathways that regulate plant defense in response to attacks by herbivorous pests (including, for example, protease inhibitors).

Increased tolerance of plants to nutrient-limited soils. Presently disclosed transcription factor genes introduced into plants may provide the means to improve uptake of essential nutrients, including nitrogenous compounds, phosphates, potassium, and trace minerals. The effect of these modifications is to increase the seedling germination and range of ornamental and crop plants. The utilities of presently disclosed transcription factor genes conferring tolerance to conditions of low nutrients also include cost savings to the grower by reducing the amounts of fertilizer needed, environmental benefits of reduced fertilizer runoff, and improved yield and stress tolerance. In addition, this gene could be used to alter seed protein amounts and/or composition that could impact yield as well as the nutritional value and production of various food products.

Hormone sensitivity. One or more of the presently disclosed transcription factor genes have been shown to affect plant abscisic acid (ABA) sensitivity. This plant hormone is likely the most important hormone in mediating the adaptation of a plant to stress. For example, ABA mediates conversion of apical meristems into dormant buds. In response to increasingly cold conditions, the newly developing leaves growing above the meristem become converted into stiff bud scales that closely wrap the meristem and protect it from mechanical damage during winter. ABA in the bud also enforces dormancy; during premature warm spells, the buds are inhibited from sprouting. Bud dormancy is eliminated after either a prolonged cold period of cold or a significant number of lengthening days. Thus, by affecting ABA sensitivity, introduced transcription factor genes may affect cold sensitivity and survivability. ABA is also important in protecting plants from drought tolerance.

Several other of the present transcription factor genes have been used to manipulate ethylene signal transduction and response pathways. These genes can thus

be used to manipulate the processes influenced by ethylene, such as seed germination or fruit ripening, and to improve seed or fruit quality.

Production of seed and leaf prenyl lipids, including tocopherol. Prenyl lipids play a role in anchoring proteins in membranes or membranous organelles. Thus modifying the prenyl lipid content of seeds and leaves could affect membrane integrity and function. A number of presently disclosed transcription factor genes have been shown to modify the tocopherol composition of plants. Tocopherols have both anti-oxidant and vitamin E activity.

Production of seed and leaf phytosterols: Presently disclosed transcription factor genes that modify levels of phytosterols in plants may have at least two utilities. First, phytosterols are an important source of precursors for the manufacture of human steroid hormones. Thus, regulation of transcription factor expression or activity could lead to elevated levels of important human steroid precursors for steroid semi-synthesis. For example, transcription factors that cause elevated levels of campesterol in leaves, or sitosterols and stigmasterols in seed crops, would be useful for this purpose. Phytosterols and their hydrogenated derivatives phytostanols also have proven cholesterol-lowering properties, and transcription factor genes that modify the expression of these compounds in plants would thus provide health benefits.

Production of seed and leaf glucosinolates. Some glucosinolates have anti-cancer activity; thus, increasing the levels or composition of these compounds by introducing several of the presently disclosed transcription factors might be of interest from a nutraceutical standpoint. (3) Glucosinolates form part of a plants natural defense against insects. Modification of glucosinolate composition or quantity could therefore afford increased protection from predators. Furthermore, in edible crops, tissue specific promoters might be used to ensure that these compounds accumulate specifically in tissues, such as the epidermis, which are not taken for consumption.

Modified seed oil content. The composition of seeds, particularly with respect to seed oil amounts and/or composition, is very important for the nutritional value and production of various food and feed products. Several of the presently disclosed transcription factor genes in seed lipid saturation that alter seed oil content could be

used to improve the heat stability of oils or to improve the nutritional quality of seed oil, by, for example, reducing the number of calories in seed, increasing the number of calories in animal feeds; or altering the ratio of saturated to unsaturated lipids comprising the oils.

Seed and leaf fatty acid composition. A number of the presently disclosed transcription factor genes have been shown to alter the fatty acid composition in plants, and seeds in particular. This modification may find particular value for improving the nutritional value of, for example, seeds or whole plants. Dietary fatty acids ratios have been shown to have an effect on, for example, bone integrity and remodeling (see, for example, Weiler, H.A., *Pediatr Res* (2000) 47:5 692-697). The ratio of dietary fatty acids may alter the precursor pools of long-chain polyunsaturated fatty acids that serve as precursors for prostaglandin synthesis. In mammalian connective tissue, prostaglandins serve as important signals regulating the balance between resorption and formation in bone and cartilage. Thus dietary fatty acid ratios altered in seeds may affect the etiology and outcome of bone loss.

Modified seed protein content. As with seed oils, the composition of seeds, particularly with respect to protein amounts and/or composition, is very important for the nutritional value and production of various food and feed products. A number of the presently disclosed transcription factor genes modify the protein concentrations in seeds would provide nutritional benefits, and may be used to prolong storage, increase seed pest or disease resistance, or modify germination rates.

Production of flavonoids in leaves and other plant parts. Expression of presently disclosed transcription factor genes that increase flavonoid production in plants, including anthocyanins and condensed tannins, may be used to alter in pigment production for horticultural purposes, and possibly increasing stress resistance. Flavonoids have antimicrobial activity and could be used to engineer pathogen resistance. Several flavonoid compounds have health promoting effects such as the inhibition of tumor growth and cancer, prevention of bone loss and the prevention of the oxidation of lipids. Increasing levels of condensed tannins, whose biosynthetic pathway is shared with anthocyanin biosynthesis, in forage legumes is an important agronomic trait because they prevent pasture bloat by collapsing protein foams within

the rumen. For a review on the utilities of flavonoids and their derivatives, refer to Dixon et al. (1999) Trends Plant Sci. 4:394-400.

Production of diterpenes in leaves and other plant parts. Depending on the plant species, varying amounts of diverse secondary biochemicals (often lipophilic terpenes) are produced and exuded or volatilized by trichomes. These exotic secondary biochemicals, which are relatively easy to extract because they are on the surface of the leaf, have been widely used in such products as flavors and aromas, drugs, pesticides and cosmetics. Thus, the overexpression of genes that are used to produce diterpenes in plants may be accomplished by introducing transcription factor genes that induce said overexpression. One class of secondary metabolites, the diterpenes, can effect several biological systems such as tumor progression, prostaglandin synthesis and tissue inflammation. In addition, diterpenes can act as insect pheromones, termite allomones, and can exhibit neurotoxic, cytotoxic and antimitotic activities. As a result of this functional diversity, diterpenes have been the target of research several pharmaceutical ventures. In most cases where the metabolic pathways are impossible to engineer, increasing trichome density or size on leaves may be the only way to increase plant productivity.

Production of anthocyanin in leaves and other plant parts. Several presently disclosed transcription factor genes can be used to alter anthocyanin production in numerous plant species. The potential utilities of these genes include alterations in pigment production for horticultural purposes, and possibly increasing stress resistance in combination with another transcription factor.

Production of miscellaneous secondary metabolites. Microarray data suggests that flux through the aromatic amino acid biosynthetic pathways and primary and secondary metabolite biosynthetic pathways are up-regulated. Presently disclosed transcription factors have been shown to be involved in regulating alkaloid biosynthesis, in part by up-regulating the enzymes indole-3-glycerol phosphatase and strictosidine synthase. Phenylalanine ammonia lyase, chalcone synthase and trans-cinnamate mono-oxygenase are also induced, and are involved in phenylpropanoid biosynthesis.



Sugar, starch, hemicellulose composition. Overexpression of the presently disclosed transcription factors that affect sugar content resulted in plants with altered leaf insoluble sugar content. Transcription factors that alter plant cell wall composition have several potential applications including altering food digestibility, plant tensile strength, wood quality, pathogen resistance and in pulp production. The potential utilities of a gene involved in glucose-specific sugar sensing are to alter energy balance, photosynthetic rate, carbohydrate accumulation, biomass production, source-sink relationships, and senescence.

Hemicellulose is not desirable in paper pulps because of its lack of strength compared with cellulose. Thus modulating the amounts of cellulose vs. hemicellulose in the plant cell wall is desirable for the paper/lumber industry. Increasing the insoluble carbohydrate content in various fruits, vegetables, and other edible consumer products will result in enhanced fiber content. Increased fiber content would not only provide health benefits in food products, but might also increase digestibility of forage crops. In addition, the hemicellulose and pectin content of fruits and berries affects the quality of jam and catsup made from them. Changes in hemicellulose and pectin content could result in a superior consumer product.

Plant response to sugars and sugar composition. In addition to their important role as an energy source and structural component of the plant cell, sugars are central regulatory molecules that control several aspects of plant physiology, metabolism and development. It is thought that this control is achieved by regulating gene expression and, in higher plants, sugars have been shown to repress or activate plant genes involved in many essential processes such as photosynthesis, glyoxylate metabolism, respiration, starch and sucrose synthesis and degradation, pathogen response, wounding response, cell cycle regulation, pigmentation, flowering and senescence. The mechanisms by which sugars control gene expression are not understood.

Because sugars are important signaling molecules, the ability to control either the concentration of a signaling sugar or how the plant perceives or responds to a signaling sugar could be used to control plant development, physiology or metabolism. For example, the flux of sucrose (a disaccharide sugar used for systemically transporting carbon and energy in most plants) has been shown to affect

gene expression and alter storage compound accumulation in seeds. Manipulation of the sucrose signaling pathway in seeds may therefore cause seeds to have more protein, oil or carbohydrate, depending on the type of manipulation. Similarly, in tubers, sucrose is converted to starch which is used as an energy store. It is thought that sugar signaling pathways may partially determine the levels of starch synthesized in the tubers. The manipulation of sugar signaling in tubers could lead to tubers with a higher starch content.

Thus, the presently disclosed transcription factor genes that manipulate the sugar signal transduction pathway may lead to altered gene expression to produce plants with desirable traits. In particular, manipulation of sugar signal transduction pathways could be used to alter source-sink relationships in seeds, tubers, roots and other storage organs leading to increase in yield.

Plant growth rate and development. A number of the presently disclosed transcription factor genes have been shown to have significant effects on plant growth rate and development. These observations have included, for example, more rapid or delayed growth and development of reproductive organs. This would provide utility for regions with short or long growing seasons, respectively. Accelerating plant growth would also improve early yield or increase biomass at an earlier stage, when such is desirable (for example, in producing forestry products).

Embryo development. Presently disclosed transcription factor genes that alter embryo development has been used to alter seed protein and oil amounts and/or composition which is very important for the nutritional value and production of various food products. Seed shape and seed coat may also be altered by these genes, which may provide for improved storage stability.

Seed germination rate. A number of the presently disclosed transcription factor genes have been shown to modify seed germination rate, including when the seeds are in conditions normally unfavorable for germination (e.g., cold, heat or salt stress, or in the presence of ABA), and may thus be used to modify and improve germination rates under adverse conditions.

Plant, seedling vigor. Seedlings transformed with presently disclosed transcription factors have been shown to possess larger cotyledons and appeared somewhat more advanced than control plants. This indicates that the seedlings developed more rapidly than the control plants. Rapid seedling development is likely to reduce loss due to diseases particularly prevalent at the seedling stage (e.g., damping off) and is thus important for survivability of plants germinating in the field or in controlled environments.

Senescence, cell death. Presently disclosed transcription factor genes may be used to alter senescence responses in plants. Although leaf senescence is thought to be an evolutionary adaptation to recycle nutrients, the ability to control senescence in an agricultural setting has significant value. For example, a delay in leaf senescence in some maize hybrids is associated with a significant increase in yields and a delay of a few days in the senescence of soybean plants can have a large impact on yield. Delayed flower senescence may also generate plants that retain their blossoms longer and this may be of potential interest to the ornamental horticulture industry.

Modified fertility. Plants that overexpress a number of the presently disclosed transcription factor genes have been shown to possess reduced fertility. This could be a desirable trait, as it could be exploited to prevent or minimize the escape of the pollen of genetically modified organisms (GMOs) into the environment.

Early and delayed flowering. Presently disclosed transcription factor genes that accelerate flowering could have valuable applications in such programs since they allow much faster generation times. In a number of species, for example, broccoli, cauliflower, where the reproductive parts of the plants constitute the crop and the vegetative tissues are discarded, it would be advantageous to accelerate time to flowering. Accelerating flowering could shorten crop and tree breeding programs. Additionally, in some instances, a faster generation time might allow additional harvests of a crop to be made within a given growing season. A number of *Arabidopsis* genes have already been shown to accelerate flowering when constitutively expressed. These include LEAFY, APETALA1 and CONSTANS (Mandel, M. et al., 1995, Nature 377, 522-524; Weigel, D. and Nilsson, O., 1995, Nature 377, 495-500; Simon et al., 1996, Nature 384, 59-62).

By regulating the expression of potential flowering using inducible promoters, flowering could be triggered by application of an inducer chemical. This would allow flowering to be synchronized across a crop and facilitate more efficient harvesting. Such inducible systems could also be used to tune the flowering of crop varieties to different latitudes. At present, species such as soybean and cotton are available as a series of maturity groups that are suitable for different latitudes on the basis of their flowering time (which is governed by day-length). A system in which flowering could be chemically controlled would allow a single high-yielding northern maturity group to be grown at any latitude. In southern regions such plants could be grown for longer, thereby increasing yields, before flowering was induced. In more northern areas, the induction would be used to ensure that the crop flowers prior to the first winter frosts.

In a sizeable number of species, for example, root crops, where the vegetative parts of the plants constitute the crop and the reproductive tissues are discarded, it would be advantageous to delay or prevent flowering. Extending vegetative development with presently disclosed transcription factor genes could thus bring about large increases in yields. Prevention of flowering might help maximize vegetative yields and prevent escape of genetically modified organism (GMO) pollen.

Extended flowering phase. Presently disclosed transcription factors that extend flowering time have utility in engineering plants with longer-lasting flowers for the horticulture industry, and for extending the time in which the plant is fertile.

Flower and leaf development. Presently disclosed transcription factor genes have been used to modify the development of flowers and leaves. This could be advantageous in the development of new ornamental cultivars that present unique configurations. In addition, some of these genes have been shown to reduce a plant's fertility, which is also useful for helping to prevent development of pollen of GMOs.

Flower abscission. Presently disclosed transcription factor genes introduced into plants have been used to retain flowers for longer periods. This would provide a significant benefit to the ornamental industry, for both cut flowers and woody plant

varieties (of, for example, maize), as well as have the potential to lengthen the fertile period of a plant, which could positively impact yield and breeding programs.

**A listing of specific effects and utilities that the presently disclosed transcription factor genes have on plants, as determined by direct observation and assay analysis, is provided in Table 4.**

#### **XVI. Antisense and Co-suppression**

In addition to expression of the nucleic acids of the invention as gene replacement or plant phenotype modification nucleic acids, the nucleic acids are also useful for sense and anti-sense suppression of expression, e.g., to down-regulate expression of a nucleic acid of the invention, e.g., as a further mechanism for modulating plant phenotype. That is, the nucleic acids of the invention, or subsequences or anti-sense sequences thereof, can be used to block expression of naturally occurring homologous nucleic acids. A variety of sense and anti-sense technologies are known in the art, e.g., as set forth in Lichtenstein and Nellen (1997) Antisense Technology: A Practical Approach IRL Press at Oxford University Press, Oxford, U.K.. In general, sense or anti-sense sequences are introduced into a cell, where they are optionally amplified, e.g., by transcription. Such sequences include both simple oligonucleotide sequences and catalytic sequences such as ribozymes.

For example, a reduction or elimination of expression (i.e., a "knock-out") of a transcription factor or transcription factor homologue polypeptide in a transgenic plant, e.g., to modify a plant trait, can be obtained by introducing an antisense construct corresponding to the polypeptide of interest as a cDNA. For antisense suppression, the transcription factor or homologue cDNA is arranged in reverse orientation (with respect to the coding sequence) relative to the promoter sequence in the expression vector. The introduced sequence need not be the full length cDNA or gene, and need not be identical to the cDNA or gene found in the plant type to be transformed. Typically, the antisense sequence need only be capable of hybridizing to the target gene or RNA of interest. Thus, where the introduced sequence is of shorter length, a higher degree of homology to the endogenous transcription factor sequence will be needed for effective antisense suppression. While antisense sequences of various lengths can be utilized, preferably, the introduced antisense sequence in the vector

will be at least 30 nucleotides in length, and improved antisense suppression will typically be observed as the length of the antisense sequence increases. Preferably, the length of the antisense sequence in the vector will be greater than 100 nucleotides. Transcription of an antisense construct as described results in the production of RNA molecules that are the reverse complement of mRNA molecules transcribed from the endogenous transcription factor gene in the plant cell.

Suppression of endogenous transcription factor gene expression can also be achieved using a ribozyme. Ribozymes are RNA molecules that possess highly specific endoribonuclease activity. The production and use of ribozymes are disclosed in U.S. Patent No. 4,987,071 and U.S. Patent No. 5,543,508. Synthetic ribozyme sequences including antisense RNAs can be used to confer RNA cleaving activity on the antisense RNA, such that endogenous mRNA molecules that hybridize to the antisense RNA are cleaved, which in turn leads to an enhanced antisense inhibition of endogenous gene expression.

Vectors in which RNA encoded by a transcription factor or transcription factor homologue cDNA is over-expressed can also be used to obtain co-suppression of a corresponding endogenous gene, e.g., in the manner described in U.S. Patent No. 5,231,020 to Jorgensen. Such co-suppression (also termed sense suppression) does not require that the entire transcription factor cDNA be introduced into the plant cells, nor does it require that the introduced sequence be exactly identical to the endogenous transcription factor gene of interest. However, as with antisense suppression, the suppressive efficiency will be enhanced as specificity of hybridization is increased, e.g., as the introduced sequence is lengthened, and/or as the sequence similarity between the introduced sequence and the endogenous transcription factor gene is increased.

Vectors expressing an untranslatable form of the transcription factor mRNA, e.g., sequences comprising one or more stop codon, or nonsense mutation) can also be used to suppress expression of an endogenous transcription factor, thereby reducing or eliminating its activity and modifying one or more traits. Methods for producing such constructs are described in U.S. Patent No. 5,583,021. Preferably, such constructs are made by introducing a premature stop codon into the transcription

factor gene. Alternatively, a plant trait can be modified by gene silencing using double-strand RNA (Sharp (1999) Genes and Development 13: 139-141). Another method for abolishing the expression of a gene is by insertion mutagenesis using the T-DNA of *Agrobacterium tumefaciens*. After generating the insertion mutants, the mutants can be screened to identify those containing the insertion in a transcription factor or transcription factor homologue gene. Plants containing a single transgene insertion event at the desired gene can be crossed to generate homozygous plants for the mutation. Such methods are well known to those of skill in the art. (See for example Koncz et al. (1992) Methods in Arabidopsis Research, World Scientific.)

Alternatively, a plant phenotype can be altered by eliminating an endogenous gene, such as a transcription factor or transcription factor homologue, e.g., by homologous recombination (Kempin et al. (1997) Nature 389:802-803).

A plant trait can also be modified by using the Cre-lox system (for example, as described in US Pat. No. 5,658,772). A plant genome can be modified to include first and second lox sites that are then contacted with a Cre recombinase. If the lox sites are in the same orientation, the intervening DNA sequence between the two sites is excised. If the lox sites are in the opposite orientation, the intervening sequence is inverted.

The polynucleotides and polypeptides of this invention can also be expressed in a plant in the absence of an expression cassette by manipulating the activity or expression level of the endogenous gene by other means. For example, by ectopically expressing a gene by T-DNA activation tagging (Ichikawa et al. (1997) Nature 390 698-701; Kakimoto et al. (1996) Science 274: 982-985). This method entails transforming a plant with a gene tag containing multiple transcriptional enhancers and once the tag has inserted into the genome, expression of a flanking gene coding sequence becomes deregulated. In another example, the transcriptional machinery in a plant can be modified so as to increase transcription levels of a polynucleotide of the invention (See, e.g., PCT Publications WO 96/06166 and WO 98/53057 which describe the modification of the DNA-binding specificity of zinc finger proteins by changing particular amino acids in the DNA-binding motif).

The transgenic plant can also include the machinery necessary for expressing or altering the activity of a polypeptide encoded by an endogenous gene, for example by altering the phosphorylation state of the polypeptide to maintain it in an activated state.

Transgenic plants (or plant cells, or plant explants, or plant tissues) incorporating the polynucleotides of the invention and/or expressing the polypeptides of the invention can be produced by a variety of well established techniques as described above. Following construction of a vector, most typically an expression cassette, including a polynucleotide, e.g., encoding a transcription factor or transcription factor homologue, of the invention, standard techniques can be used to introduce the polynucleotide into a plant, a plant cell, a plant explant or a plant tissue of interest. Optionally, the plant cell, explant or tissue can be regenerated to produce a transgenic plant.

The plant can be any higher plant, including gymnosperms, monocotyledonous and dicotyledonous plants. Suitable protocols are available for *Leguminosae* (alfalfa, soybean, clover, etc.), *Umbelliferae* (carrot, celery, parsnip), *Cruciferae* (cabbage, radish, rapeseed, broccoli, etc.), *Curcubitaceae* (melons and cucumber), *Gramineae* (wheat, corn, rice, barley, millet, etc.), *Solanaceae* (potato, tomato, tobacco, peppers, etc.), and various other crops. See protocols described in Ammirato et al. (1984) Handbook of Plant Cell Culture—Crop Species, Macmillan Publ. Co. Shimamoto et al. (1989) Nature 338:274-276; Fromm et al. (1990) Bio/Technology 8:833-839; and Vasil et al. (1990) Bio/Technology 8:429-434.

Transformation and regeneration of both monocotyledonous and dicotyledonous plant cells is now routine, and the selection of the most appropriate transformation technique will be determined by the practitioner. The choice of method will vary with the type of plant to be transformed; those skilled in the art will recognize the suitability of particular methods for given plant types. Suitable methods can include, but are not limited to: electroporation of plant protoplasts; liposome-mediated transformation; polyethylene glycol (PEG) mediated transformation; transformation using viruses; micro-injection of plant cells; micro-projectile bombardment of plant cells; vacuum infiltration; and *Agrobacterium tumefaciens* mediated transformation.



Transformation means introducing a nucleotide sequence into a plant in a manner to cause stable or transient expression of the sequence.

Successful examples of the modification of plant characteristics by transformation with cloned sequences which serve to illustrate the current knowledge in this field of technology, and which are herein incorporated by reference, include: U.S. Patent Nos. 5,571,706; 5,677,175; 5,510,471; 5,750,386; 5,597,945; 5,589,615; 5,750,871; 5,268,526; 5,780,708; 5,538,880; 5,773,269; 5,736,369 and 5,610,042.

Following transformation, plants are preferably selected using a dominant selectable marker incorporated into the transformation vector. Typically, such a marker will confer antibiotic or herbicide resistance on the transformed plants, and selection of transformants can be accomplished by exposing the plants to appropriate concentrations of the antibiotic or herbicide.

After transformed plants are selected and grown to maturity, those plants showing a modified trait are identified. The modified trait can be any of those traits described above. Additionally, to confirm that the modified trait is due to changes in expression levels or activity of the polypeptide or polynucleotide of the invention can be determined by analyzing mRNA expression using Northern blots, RT-PCR or microarrays, or protein expression using immunoblots or Western blots or gel shift assays.

#### **XVII. Integrated Systems – Sequence Identity**

Additionally, the present invention may be an integrated system, computer or computer readable medium that comprises an instruction set for determining the identity of one or more sequences in a database. In addition, the instruction set can be used to generate or identify sequences that meet any specified criteria. Furthermore, the instruction set may be used to associate or link certain functional benefits, such improved characteristics, with one or more identified sequence.

For example, the instruction set can include, e.g., a sequence comparison or other alignment program, e.g., an available program such as, for example, the Wisconsin Package Version 10.0, such as BLAST, FASTA, PILEUP,

FINDPATTERNS or the like (GCG, Madison, WI). Public sequence databases such as GenBank, EMBL, Swiss-Prot and PIR or private sequence databases such as PHYTOSEQ sequence database (Incyte Genomics, Palo Alto, CA) can be searched.

Alignment of sequences for comparison can be conducted by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443-453, by the search for similarity method of Pearson and Lipman (1988) Proc. Natl. Acad. Sci. U.S.A. 85:2444-2448, by computerized implementations of these algorithms. After alignment, sequence comparisons between two (or more) polynucleotides or polypeptides are typically performed by comparing sequences of the two sequences over a comparison window to identify and compare local regions of sequence similarity. The comparison window can be a segment of at least about 20 contiguous positions, usually about 50 to about 200, more usually about 100 to about 150 contiguous positions. A description of the method is provided in Ausubel et al., *supra*.

A variety of methods for determining sequence relationships can be used, including manual alignment and computer assisted sequence alignment and analysis. This later approach is a preferred approach in the present invention, due to the increased throughput afforded by computer assisted methods. As noted above, a variety of computer programs for performing sequence alignment are available, or can be produced by one of skill.

One example algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul et al. J. Mol. Biol 215:403-410 (1990). Software for performing BLAST analyses is publicly available, e.g., through the National Center for Biotechnology Information (see internet website at [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov)). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to

find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always  $> 0$ ) and N (penalty score for mismatching residues; always  $< 0$ ). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915). Unless otherwise indicated, "sequence identity" here refers to the % sequence identity generated from a tblastx using the NCBI version of the algorithm at the default settings using gapped alignments with the filter "off" (*see*, for example, internet website at [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov)).

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see*, e.g., Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence (and, therefore, in this context, homologous) if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, or less than about 0.01, and or even less than about 0.001. An additional example of a useful sequence alignment algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. The program can align, e.g., up to 300 sequences of a maximum length of 5,000 letters.

The integrated system, or computer typically includes a user input interface allowing a user to selectively view one or more sequence records corresponding to the one or more character strings, as well as an instruction set which aligns the one or more character strings with each other or with an additional character string to identify one or more region of sequence similarity. The system may include a link of one or more character strings with a particular phenotype or gene function. Typically, the system includes a user readable output element that displays an alignment produced by the alignment instruction set.

The methods of this invention can be implemented in a localized or distributed computing environment. In a distributed environment, the methods may implemented on a single computer comprising multiple processors or on a multiplicity of computers. The computers can be linked, e.g. through a common bus, but more preferably the computer(s) are nodes on a network. The network can be a generalized or a dedicated local or wide-area network and, in certain preferred embodiments, the computers may be components of an intra-net or an internet.

Thus, the invention provides methods for identifying a sequence similar or homologous to one or more polynucleotides as noted herein, or one or more target polypeptides encoded by the polynucleotides, or otherwise noted herein and may include linking or associating a given plant phenotype or gene function with a sequence. In the methods, a sequence database is provided (locally or across an inter or intra net) and a query is made against the sequence database using the relevant sequences herein and associated plant phenotypes or gene functions.

Any sequence herein can be entered into the database, before or after querying the database. This provides for both expansion of the database and, if done before the querying step, for insertion of control sequences into the database. The control sequences can be detected by the query to ensure the general integrity of both the database and the query. As noted, the query can be performed using a web browser based interface. For example, the database can be a centralized public database such as those noted herein, and the querying can be done from a remote terminal or computer across an internet or intranet.

### **XVIII. Examples**

The following examples are intended to illustrate but not limit the present invention. The complete descriptions of the traits associated with each polynucleotide of the invention is fully disclosed in Table 4, Table 5, and Table 6.

#### **Example I: Full Length Gene Identification and Cloning**

Putative transcription factor sequences (genomic or ESTs) related to known transcription factors were identified in the *Arabidopsis thaliana* GenBank database using the tblastn sequence analysis program using default parameters and a P-value cutoff threshold of -4 or -5 or lower, depending on the length of the query sequence. Putative transcription factor sequence hits were then screened to identify those containing particular sequence strings. If the sequence hits contained such sequence strings, the sequences were confirmed as transcription factors.

Alternatively, *Arabidopsis thaliana* cDNA libraries derived from different tissues or treatments, or genomic libraries were screened to identify novel members of a transcription family using a low stringency hybridization approach. Probes were synthesized using gene specific primers in a standard PCR reaction (annealing temperature 60° C) and labeled with <sup>32</sup>P dCTP using the High Prime DNA Labeling Kit (Boehringer Mannheim). Purified radiolabelled probes were added to filters immersed in Church hybridization medium (0.5 M NaPO<sub>4</sub> pH 7.0, 7% SDS, 1 % w/v bovine serum albumin) and hybridized overnight at 60°C with shaking. Filters were washed two times for 45 to 60 minutes with 1xSSC, 1% SDS at 60° C.

To identify additional sequence 5' or 3' of a partial cDNA sequence in a cDNA library, 5' and 3' rapid amplification of cDNA ends (RACE) was performed using the Marathon<sup>TM</sup> cDNA amplification kit (Clontech, Palo Alto, CA). Generally, the method entailed first isolating poly(A) mRNA, performing first and second strand cDNA synthesis to generate double stranded cDNA, blunting cDNA ends, followed by ligation of the Marathon<sup>TM</sup> Adaptor to the cDNA to form a library of adaptor-ligated ds cDNA.

Gene-specific primers were designed to be used along with adaptor specific primers for both 5' and 3' RACE reactions. Nested primers, rather than single

primers, were used to increase PCR specificity. Using 5' and 3' RACE reactions, 5' and 3' RACE fragments were obtained, sequenced and cloned. The process can be repeated until 5' and 3' ends of the full-length gene were identified. Then the full-length cDNA was generated by PCR using primers specific to 5' and 3' ends of the gene by end-to-end PCR.

#### **Example II: Construction of Expression Vectors**

The sequence was amplified from a genomic or cDNA library using primers specific to sequences upstream and downstream of the coding region. The expression vector was pMEN20 or pMEN65, which are both derived from pMON316 (Sanders et al, (1987) Nucleic Acids Research 15:1543-1558) and contain the CaMV 35S promoter to express transgenes. To clone the sequence into the vector, both pMEN20 and the amplified DNA fragment were digested separately with SalI and NotI restriction enzymes at 37° C for 2 hours. The digestion products were subject to electrophoresis in a 0.8% agarose gel and visualized by ethidium bromide staining. The DNA fragments containing the sequence and the linearized plasmid were excised and purified by using a Qiaquick gel extraction kit (Qiagen, Valencia CA). The fragments of interest were ligated at a ratio of 3:1 (vector to insert). Ligation reactions using T4 DNA ligase (New England Biolabs, Beverly MA) were carried out at 16° C for 16 hours. The ligated DNAs were transformed into competent cells of the *E. coli* strain DH5alpha by using the heat shock method. The transformations were plated on LB plates containing 50 mg/l kanamycin (Sigma, St. Louis, MO). Individual colonies were grown overnight in five milliliters of LB broth containing 50 mg/l kanamycin at 37° C. Plasmid DNA was purified by using Qiaquick Mini Prep kits (Qiagen).

#### **Example III: Transformation of *Agrobacterium* with the Expression Vector**

After the plasmid vector containing the gene was constructed, the vector was used to transform *Agrobacterium tumefaciens* cells expressing the gene products. The stock of *Agrobacterium tumefaciens* cells for transformation were made as described by Nagel et al. (1990) FEMS Microbiol Letts. 67: 325-328. *Agrobacterium* strain ABI was grown in 250 ml LB medium (Sigma) overnight at 28°C with shaking until an absorbance ( $A_{600}$ ) of 0.5 – 1.0 was reached. Cells were harvested by centrifugation at 4,000 x g for 15 min at 4° C. Cells were then resuspended in 250 µl chilled buffer

(1 mM HEPES, pH adjusted to 7.0 with KOH). Cells were centrifuged again as described above and resuspended in 125  $\mu$ l chilled buffer. Cells were then centrifuged and resuspended two more times in the same HEPES buffer as described above at a volume of 100  $\mu$ l and 750  $\mu$ l, respectively. Resuspended cells were then distributed into 40  $\mu$ l aliquots, quickly frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$ .

*Agrobacterium* cells were transformed with plasmids prepared as described above following the protocol described by Nagel et al. For each DNA construct to be transformed, 50 – 100 ng DNA (generally resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0) was mixed with 40  $\mu$ l of *Agrobacterium* cells. The DNA/cell mixture was then transferred to a chilled cuvette with a 2mm electrode gap and subject to a 2.5 kV charge dissipated at 25  $\mu$ F and 200  $\mu$ F using a Gene Pulser II apparatus (Bio-Rad, Hercules, CA). After electroporation, cells were immediately resuspended in 1.0 ml LB and allowed to recover without antibiotic selection for 2 – 4 hours at  $28^{\circ}\text{C}$  in a shaking incubator. After recovery, cells were plated onto selective medium of LB broth containing 100  $\mu$ g/ml spectinomycin (Sigma) and incubated for 24-48 hours at  $28^{\circ}\text{C}$ . Single colonies were then picked and inoculated in fresh medium. The presence of the plasmid construct was verified by PCR amplification and sequence analysis.

#### **Example IV: Transformation of *Arabidopsis* Plants with *Agrobacterium tumefaciens* with Expression Vector**

After transformation of *Agrobacterium tumefaciens* with plasmid vectors containing the gene, single *Agrobacterium* colonies were identified, propagated, and used to transform *Arabidopsis* plants. Briefly, 500 ml cultures of LB medium containing 50 mg/l kanamycin were inoculated with the colonies and grown at  $28^{\circ}\text{C}$  with shaking for 2 days until an optical absorbance at 600 nm wavelength over 1 cm ( $A_{600}$ ) of  $> 2.0$  is reached. Cells were then harvested by centrifugation at 4,000 x g for 10 min, and resuspended in infiltration medium (1/2 X Murashige and Skoog salts (Sigma), 1 X Gamborg's B-5 vitamins (Sigma), 5.0% (w/v) sucrose (Sigma), 0.044  $\mu$ M benzylamino purine (Sigma), 200  $\mu$ l/l Silwet L-77 (Lehle Seeds) until an  $A_{600}$  of 0.8 was reached.

Prior to transformation, *Arabidopsis thaliana* seeds (ecotype Columbia) were sown at a density of ~10 plants per 4" pot onto Pro-Mix BX potting medium (Hummert International) covered with fiberglass mesh (18 mm X 16 mm). Plants were grown under continuous illumination (50-75  $\mu\text{E}/\text{m}^2/\text{sec}$ ) at 22-23° C with 65-70% relative humidity. After about 4 weeks, primary inflorescence stems (bolts) are cut off to encourage growth of multiple secondary bolts. After flowering of the mature secondary bolts, plants were prepared for transformation by removal of all siliques and opened flowers.

The pots were then immersed upside down in the mixture of *Agrobacterium* infiltration medium as described above for 30 sec, and placed on their sides to allow draining into a 1' x 2' flat surface covered with plastic wrap. After 24 h, the plastic wrap was removed and pots are turned upright. The immersion procedure was repeated one week later, for a total of two immersions per pot. Seeds were then collected from each transformation pot and analyzed following the protocol described below.

#### **Example V: Identification of *Arabidopsis* Primary Transformants**

Seeds collected from the transformation pots were sterilized essentially as follows. Seeds were dispersed into in a solution containing 0.1% (v/v) Triton X-100 (Sigma) and sterile H<sub>2</sub>O and washed by shaking the suspension for 20 min. The wash solution was then drained and replaced with fresh wash solution to wash the seeds for 20 min with shaking. After removal of the second wash solution, a solution containing 0.1% (v/v) Triton X-100 and 70% ethanol (Equistar) was added to the seeds and the suspension was shaken for 5 min. After removal of the ethanol/detergent solution, a solution containing 0.1% (v/v) Triton X-100 and 30% (v/v) bleach (Clorox) was added to the seeds, and the suspension was shaken for 10 min. After removal of the bleach/detergent solution, seeds were then washed five times in sterile distilled H<sub>2</sub>O. The seeds were stored in the last wash water at 4° C for 2 days in the dark before being plated onto antibiotic selection medium (1 X Murashige and Skoog salts (pH adjusted to 5.7 with 1M KOH), 1 X Gamborg's B-5 vitamins, 0.9% phytagar (Life Technologies), and 50 mg/l kanamycin). Seeds were germinated under continuous illumination (50-75  $\mu\text{E}/\text{m}^2/\text{sec}$ ) at 22-23° C. After 7-10 days of growth under these conditions, kanamycin resistant primary transformants (T<sub>1</sub>



generation) were visible and obtained. These seedlings were transferred first to fresh selection plates where the seedlings continued to grow for 3-5 more days, and then to soil (Pro-Mix BX potting medium).

Primary transformants were crossed and progeny seeds ( $T_2$ ) collected; kanamycin resistant seedlings were selected and analyzed. The expression levels of the recombinant polynucleotides in the transformants varies from about a 5% expression level increase to a least a 100% expression level increase. Similar observations are made with respect to polypeptide level expression.

#### **Example VI: Identification of *Arabidopsis* Plants with Transcription Factor Gene Knockouts**

The screening of insertion mutagenized *Arabidopsis* collections for null mutants in a known target gene was essentially as described in Krysan et al (1999) Plant Cell 11:2283-2290. Briefly, gene-specific primers, nested by 5-250 base pairs to each other, were designed from the 5' and 3' regions of a known target gene. Similarly, nested sets of primers were also created specific to each of the T-DNA or transposon ends (the "right" and "left" borders). All possible combinations of gene specific and T-DNA/transposon primers were used to detect by PCR an insertion event within or close to the target gene. The amplified DNA fragments were then sequenced which allows the precise determination of the T-DNA/transposon insertion point relative to the target gene. Insertion events within the coding or intervening sequence of the genes were deconvoluted from a pool comprising a plurality of insertion events to a single unique mutant plant for functional characterization. The method is described in more detail in Yu and Adam, US Application Serial No. 09/177,733 filed October 23, 1998.

#### **Example VII: Identification of Modified Phenotypes in Overexpression or Gene Knockout Plants**

Experiments were performed to identify those transformants or knockouts that exhibited modified biochemical characteristics. Among the biochemicals that were assayed were insoluble sugars, such as arabinose, fucose, galactose, mannose, rhamnose or xylose or the like; prenol lipids, such as lutein, beta-carotene, xanthophyll-1, xanthophyll-2, chlorophylls A or B, or alpha-, delta- or gamma-

tocopherol or the like; fatty acids, such as 16:0 (palmitic acid), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (linoleic acid), 20:0, 18:3 (linolenic acid), 20:1 (eicosenoic acid), 20:2, 22:1 (erucic acid) or the like; waxes, such as by altering the levels of C29, C31, or C33 alkanes; sterols, such as brassicasterol, campesterol, stigmasterol, sitosterol or stigmastanol or the like, glucosinolates, protein or oil levels.

Fatty acids were measured using two methods depending on whether the tissue was from leaves or seeds. For leaves, lipids were extracted and esterified with hot methanolic  $\text{H}_2\text{SO}_4$  and partitioned into hexane from methanolic brine. For seed fatty acids, seeds were pulverized and extracted in methanol:heptane:toluene:2,2-dimethoxypropane: $\text{H}_2\text{SO}_4$  (39:34:20:5:2) for 90 minutes at 80°C. After cooling to room temperature the upper phase, containing the seed fatty acid esters, was subjected to GC analysis. Fatty acid esters from both seed and leaf tissues were analyzed with a Supelco SP-2330 column.

Glucosinolates were purified from seeds or leaves by first heating the tissue at 95°C for 10 minutes. Preheated ethanol:water (50:50) is and after heating at 95°C for a further 10 minutes, the extraction solvent is applied to a DEAE Sephadex column which had been previously equilibrated with 0.5 M pyridine acetate. Desulfoglucosinolates were eluted with 300  $\mu\text{l}$  water and analyzed by reverse phase HPLC monitoring at 226 nm.

For wax alkanes, samples were extracted using an identical method as fatty acids and extracts were analyzed on a HP 5890 GC coupled with a 5973 MSD. Samples were chromatographically isolated on a J&W DB35 mass spectrometer (J&W Scientific).

To measure prenyl lipids levels, seeds or leaves were pulverized with 1 to 2% pyrogallol as an antioxidant. For seeds, extracted samples were filtered and a portion removed for tocopherol and carotenoid/chlorophyll analysis by HPLC. The remaining material was saponified for sterol determination. For leaves, an aliquot was removed and diluted with methanol and chlorophyll A, chlorophyll B, and total carotenoids measured by spectrophotometry by determining optical absorbance at

665.2 nm, 652.5 nm, and 470 nm. An aliquot was removed for tocopherol and carotenoid/chlorophyll composition by HPLC using a Waters uBondapak C18 column (4.6 mm x 150 mm). The remaining methanolic solution was saponified with 10% KOH at 80°C for one hour. The samples were cooled and diluted with a mixture of methanol and water. A solution of 2% methylene chloride in hexane was mixed in and the samples were centrifuged. The aqueous methanol phase was again re-extracted 2% methylene chloride in hexane and, after centrifugation, the two upper phases were combined and evaporated. 2% methylene chloride in hexane was added to the tubes and the samples were then extracted with one ml of water. The upper phase was removed, dried, and resuspended in 400  $\mu$ l of 2% methylene chloride in hexane and analyzed by gas chromatography using a 50 m DB-5ms (0.25 mm ID, 0.25  $\mu$ m phase, J&W Scientific).

Suppression of endogenous transcription factor gene expression can also be achieved using RNA interference, or RNAi. RNAi is a post-transcriptional, targeted gene-silencing technique that uses double-stranded RNA (dsRNA) to incite degradation of messenger RNA (mRNA) containing the same sequence as the dsRNA (Constans, (2002) *The Scientist* 16:36). Small interfering RNAs, or siRNAs are produced in at least two steps: an endogenous ribonuclease cleaves longer dsRNA into shorter, 21-23 nucleotide-long RNAs. The siRNA segments then mediate the degradation of the target mRNA (Zamore, (2001) *Nature Struct. Biol.*, 8:746-50). RNAi has been used for gene function determination in a manner similar to antisense oligonucleotides (Constans, (2002) *The Scientist* 16:36). Expression vectors that continually express siRNAs in transiently and stably transfected have been engineered to express small hairpin RNAs (shRNAs), which get processed in vivo into siRNA-like molecules capable of carrying out gene-specific silencing (Brummelkamp et al., (2002) *Science* 296:550-553, and Paddison, et al. (2002) *Genes & Dev.* 16:948-958). Post-transcriptional gene silencing by double-stranded RNA is discussed in further detail by Hammond et al. (2001) *Nature Rev Gen* 2: 110-119, Fire et al. (1998) *Nature* 391: 806-811 and Timmons and Fire (1998) *Nature* 395: 854.--

Insoluble sugar levels were measured by the method essentially described by Reiter et al., *Plant Journal* 12:335-345. This method analyzes the neutral sugar composition of cell wall polymers found in *Arabidopsis* leaves. Soluble sugars were

separated from sugar polymers by extracting leaves with hot 70% ethanol. The remaining residue containing the insoluble polysaccharides was then acid hydrolyzed with allose added as an internal standard. Sugar monomers generated by the hydrolysis were then reduced to the corresponding alditols by treatment with NaBH<sub>4</sub>, then were acetylated to generate the volatile alditol acetates which were then analyzed by GC-FID. Identity of the peaks was determined by comparing the retention times of known sugars converted to the corresponding alditol acetates with the retention times of peaks from wild-type plant extracts. Alditol acetates were analyzed on a Supelco SP-2330 capillary column (30 m x 250  $\mu$ m x 0.2  $\mu$ m) using a temperature program beginning at 180° C for 2 minutes followed by an increase to 220° C in 4 minutes. After holding at 220° C for 10 minutes, the oven temperature is increased to 240° C in 2 minutes and held at this temperature for 10 minutes and brought back to room temperature.

To identify plants with alterations in total seed oil or protein content, 150mg of seeds from T2 progeny plants were subjected to analysis by Near Infrared Reflectance Spectroscopy (NIRS) using a Foss NirSystems Model 6500 with a spinning cup transport system. NIRS is a non-destructive analytical method used to determine seed oil and protein composition. Infrared is the region of the electromagnetic spectrum located after the visible region in the direction of longer wavelengths. 'Near infrared' owns its name for being the infrared region near to the visible region of the electromagnetic spectrum. For practical purposes, near infrared comprises wavelengths between 800 and 2500 nm. NIRS is applied to organic compounds rich in O-H bonds (such as moisture, carbohydrates, and fats), C-H bonds (such as organic compounds and petroleum derivatives), and N-H bonds (such as proteins and amino acids). The NIRS analytical instruments operate by statistically correlating NIRS signals at several wavelengths with the characteristic or property intended to be measured. All biological substances contain thousands of C-H, O-H, and N-H bonds. Therefore, the exposure to near infrared radiation of a biological sample, such as a seed, results in a complex spectrum which contains qualitative and quantitative information about the physical and chemical composition of that sample.

The numerical value of a specific analyte in the sample, such as protein content or oil content, is mediated by a calibration approach known as chemometrics.

Chemometrics applies statistical methods such as multiple linear regression (MLR), partial least squares (PLS), and principle component analysis (PCA) to the spectral data and correlates them with a physical property or other factor, that property or factor is directly determined rather than the analyte concentration itself. The method first provides "wet chemistry" data of the samples required to develop the calibration.

Calibration for Arabidopsis seed oil composition was performed using accelerated solvent extraction using 1 g seed sample size and was validated against certified canola seed. A similar wet chemistry approach was performed for seed protein composition calibration.

Data obtained from NIRS analysis was analyzed statistically using a nearest-neighbor (N-N) analysis. The N-N analysis allows removal of within-block spatial variability in a fairly flexible fashion which does not require prior knowledge of the pattern of variability in the chamber. Ideally, all hybrids are grown under identical experimental conditions within a block (rep). In reality, even in many block designs, significant within-block variability exists. Nearest-neighbor procedures are based on assumption that environmental effect of a plot is closely related to that of its neighbors. Nearest-neighbor methods use information from adjacent plots to adjust for within-block heterogeneity and so provide more precise estimates of treatment means and differences. If there is within-plot heterogeneity on a spatial scale that is larger than a single plot and smaller than the entire block, then yields from adjacent plots will be positively correlated. Information from neighboring plots can be used to reduce or remove the unwanted effect of the spatial heterogeneity, and hence improve the estimate of the treatment effect. Data from neighboring plots can also be used to reduce the influence of competition between adjacent plots. The Papadakis N-N analysis can be used with designs to remove within-block variability that would not be removed with the standard split plot analysis (Papadakis, 1973, Inst. d'Amelior. Plantes Thessaloniki (Greece) Bull. Scientif., No. 23; Papadakis, 1984, Proc. Acad. Athens, 59, 326-342).

Experiments were performed to identify those transformants or knockouts that exhibited an improved pathogen tolerance. For such studies, the transformants were exposed to biotrophic fungal pathogens, such as *Erysiphe orontii*, and necrotropic

fungus pathogens, such as *Fusarium oxysporum*. *Fusarium oxysporum* isolates cause vascular wilts and damping off of various annual vegetables, perennials and weeds (Mauch-Mani and Slusarenko (1994) Molecular Plant-Microbe Interactions 7: 378-383). For *Fusarium oxysporum* experiments, plants grown on Petri dishes were sprayed with a fresh spore suspension of *F. oxysporum*. The spore suspension was prepared as follows: A plug of fungal hyphae from a plate culture was placed on a fresh potato dextrose agar plate and allowed to spread for one week. 5 ml sterile water was then added to the plate, swirled, and pipetted into 50 ml Armstrong *Fusarium* medium. Spores were grown overnight in *Fusarium* medium and then sprayed onto plants using a Preval paint sprayer. Plant tissue was harvested and frozen in liquid nitrogen 48 hours post infection.

*Erysiphe orontii* is a causal agent of powdery mildew. For *Erysiphe orontii* experiments, plants were grown approximately 4 weeks in a greenhouse under 12 hour light (20°C, ~30% relative humidity (rh)). Individual leaves were infected with *E. orontii* spores from infected plants using a camel's hair brush, and the plants were transferred to a Percival growth chamber (20°C, 80% rh.). Plant tissue was harvested and frozen in liquid nitrogen 7 days post infection.

*Botrytis cinerea* is a necrotrophic pathogen. *Botrytis cinerea* was grown on potato dextrose agar in the light. A spore culture was made by spreading 10 ml of sterile water on the fungus plate, swirling and transferring spores to 10 ml of sterile water. The spore inoculum (approx. 105 spores/ml) was used to spray 10 day-old seedlings grown under sterile conditions on MS (minus sucrose) media. Symptoms were evaluated every day up to approximately 1 week.

Infection with bacterial pathogens *Pseudomonas syringae* pv *maculicola* (Psm) strain 4326 and pv *maculicola* strain 4326 was performed by hand inoculation at two doses. Two inoculation doses allows the differentiation between plants with enhanced susceptibility and plants with enhanced resistance to the pathogen. Plants were grown for 3 weeks in the greenhouse, then transferred to the growth chamber for the remainder of their growth. Psm ES4326 was hand inoculated with 1 ml syringe on 3 fully-expanded leaves per plant (4 1/2 wk old), using at least 9 plants per overexpressing line at two inoculation doses, OD=0.005 and OD=0.0005. Disease

scoring occurred at day 3 post-inoculation with pictures of the plants and leaves taken in parallel.

In some instances, expression patterns of the pathogen-induced genes (such as defense genes) was monitored by microarray experiments. cDNAs were generated by PCR and resuspended at a final concentration of ~ 100 ng/ul in 3X SSC or 150mM Na-phosphate (Eisen and Brown (1999) *Methods Enzymol.* 303:179-205). The cDNAs were spotted on microscope glass slides coated with polylysine. The prepared cDNAs were aliquoted into 384 well plates and spotted on the slides using an x-y-z gantry (OmniGrid) purchased from GeneMachines (Menlo Park, CA) outfitted with quill type pins purchased from Telechem International (Sunnyvale, CA). After spotting, the arrays were cured for a minimum of one week at room temperature, rehydrated and blocked following the protocol recommended by Eisen and Brown (1999; supra).

Sample total RNA (10 ug) samples were labeled using fluorescent Cy3 and Cy5 dyes. Labeled samples were resuspended in 4X SSC/0.03% SDS/4 ug salmon sperm DNA/2 ug tRNA/ 50mM Na-pyrophosphate, heated for 95°C for 2.5 minutes, spun down and placed on the array. The array was then covered with a glass coverslip and placed in a sealed chamber. The chamber was then kept in a water bath at 62°C overnight. The arrays were washed as described in Eisen and Brown (1999) and scanned on a General Scanning 3000 laser scanner. The resulting files are subsequently quantified using Imagene, a software purchased from BioDiscovery (Los Angeles, CA).

Experiments were performed to identify those transformants or knockouts that exhibited an improved environmental stress tolerance. For such studies, the transformants were exposed to a variety of environmental stresses. Plants were exposed to chilling stress (6 hour exposure to 4-8°C), heat stress (6 hour exposure to 32-37°C), high salt stress (6 hour exposure to 200 mM NaCl), drought stress (168 hours after removing water from trays), osmotic stress (6 hour exposure to 3 M mannitol), or nutrient limitation (nitrogen, phosphate, and potassium) (Nitrogen: all components of MS medium remained constant except N was reduced to 20 mg/l of  $\text{NH}_4\text{NO}_3$ , or Phosphate: All components of MS medium except  $\text{KH}_2\text{PO}_4$ , which was

replaced by  $K_2SO_4$ , Potassium: All components of MS medium except removal of  $KNO_3$  and  $KH_2PO_4$ , which were replaced by  $NaH_4PO_4$ ).

Experiments were performed to identify those transformants or knockouts that exhibited a modified structure and development characteristics. For such studies, the transformants were observed by eye to identify novel structural or developmental characteristics associated with the ectopic expression of the polynucleotides or polypeptides of the invention.

Experiments were performed to identify those transformants or knockouts that exhibited modified sugar-sensing. For such studies, seeds from transformants were germinated on media containing 5% glucose or 9.4% sucrose which normally partially restrict hypocotyl elongation. Plants with altered sugar sensing may have either longer or shorter hypocotyls than normal plants when grown on this media. Additionally, other plant traits may be varied such as root mass.

Flowering time was measured by the number of rosette leaves present when a visible inflorescence of approximately 3 cm is apparent. Rosette and total leaf number on the progeny stem are tightly correlated with the timing of flowering (Koornneef et al (1991) *Mol. Gen. Genet* 229:57-66). The vernalization response was measured. For vernalization treatments, seeds were sown to MS agar plates, sealed with micropore tape, and placed in a 4°C cold room with low light levels for 6-8 weeks. The plates were then transferred to the growth rooms alongside plates containing freshly sown non-vernalized controls. Rosette leaves were counted when a visible inflorescence of approximately 3 cm was apparent.

Modified phenotypes observed for particular overexpressor or knockout plants are provided in Table 4. For a particular overexpressor that shows a less beneficial characteristic, it may be more useful to select a plant with a decreased expression of the particular transcription factor. For a particular knockout that shows a less beneficial characteristic, it may be more useful to select a plant with an increased expression of the particular transcription factor.



The sequences of the Sequence Listing or those in Tables 4 or 5 or those disclosed here can be used to prepare transgenic plants and plants with altered traits. The specific transgenic plants listed below are produced from the sequences of the Sequence Listing, as noted. Table 4 provides exemplary polynucleotide and polypeptide sequences of the invention. Table 4 includes, from left to right for each sequence: the first column shows the polynucleotide SEQ ID NO; the second column shows the Mendel Gene ID No., GID; the third column shows the trait(s) resulting from the knock out or overexpression of the polynucleotide in the transgenic plant; the fourth column shows the category of the trait; the fifth column shows the transcription factor family to which the polynucleotide belongs; the sixth column ("Comment"), includes specific effects and utilities conferred by the polynucleotide of the first column; the seventh column shows the SEQ ID NO of the polypeptide encoded by the polynucleotide; and the eighth column shows the amino acid residue positions of the conserved domain in amino acid (AA) co-ordinates.

G720: The complete sequence of G720 (SEQ ID NO: 25); similar to a portion of APRR2, Arabidopsis pseudo-response regulator (APRR2; Makino et al. 2000 Plant Cell Physiol. 41:791-803) was determined. A line homozygous for a T-DNA insertion in G720 and lines overexpressing G720 under the 35S promoter were used to determine the function of this gene. The T-DNA insertion in G720 was approximately half-way into the coding sequence, just before the conserved domain, and therefore should result in a null mutation. G720 knockout mutants were slightly more sensitive to freezing than the wild-type controls when the seedlings were cold-acclimated prior to freezing. G720 overexpressing lines were slightly more tolerant to freezing. When seedlings were frozen at  $-10^{\circ}\text{C}$  for 20 hours, the G720 plants recovered slightly better compared to the wild-type control in two separate experiments. G720 was induced by ABA, salt, osmotic stress, drought, heat, and auxin. The combination of enhanced sensitivity to freezing in the knockout mutants, enhanced resistance in the overexpressing lines, and the induction pattern of G720 comprised strong evidence that G720 functions in regulation of dehydration tolerance, as freezing is a form of dehydration stress.

Plants overexpressing G720 also showed reduced time to flowering in the T1 generation. One third of the 35S::G720 T1 seedlings, from each of two separate

batches, flowered markedly earlier (up to 1 week sooner, 24-hour light conditions) than controls plants. All of the T1 lines showed high levels of G720 overexpression (determined by RT-PCR). Three early flowering T1 plants were selected for further study. However, none of these lines flowered early in the T2, suggesting that activity of the transgene might have been reduced between the generations.

#### Closely Related Genes from Other Species

G720 showed significant similarity to a drought-induced *M. truncatula* EST, GenBank accession number BG450227, that encodes a pseudo-receiver domain. The sequence similarity is high enough to suggest that the two proteins are orthologs, and the fact that G720 was also drought-induced is consistent with this hypothesis. Other ESTs from tomato and potato (BG642566, BG128919, BG129142, and BG887673) also showed high similarity to G720 and represent potential orthologs.

G1792: G1792 (SEQ ID NO: 45) was studied using transgenic plants in which the gene was expressed under the control of the 35S promoter. 35S::G1792 plants were more tolerant to the fungal pathogens *Fusarium oxysporum* and *Botrytis cinerea* and showed fewer symptoms after inoculation with a low dose of each pathogen. This result was confirmed using individual T2 lines. The effect of G1792 overexpression in increasing tolerance to pathogens received further, incidental confirmation. T2 plants of 35S::G1792 lines 5 and 12 had been growing in a room that suffered a serious powdery mildew infection. For each line, a pot of 6 plants was present in a flat containing 9 other pots of lines from unrelated genes. In either of the two different flats, the only plants that were free from infection were those from the 35S::G1792 line. This observation suggests that G1792 overexpression might increase resistance to powdery mildew. Additional experiments confirmed that 35S::G1792 plants showed increased tolerance to Erysiphe. G1792 was ubiquitously expressed, but appears to be induced by salicylic acid.

35S::G1792 overexpressing plants also showed more tolerance to growth under nitrogen-limiting conditions. In a root growth assay under conditions of limiting N, 35S::G1792 lines were slightly less stunted. In a germination assay that monitors the effect of C on N signaling through anthocyanin production on high sucrose plus and

minus glutamine the 35S::G1792 lines make less anthocyanin on high sucrose plus glutamine, suggesting that the gene could be involved in the plants ability to monitor their carbon and nitrogen status.

G1792 overexpressing plants showed several mild morphological alterations: leaves were dark green and shiny, and plants bolted, subsequently senesced, slightly later than wild-type controls. Among the T1 plants, additional morphological variation (not reproduced later in the T2 plants) was observed: many showed reductions in size as well as aberrations in leaf shape, phyllotaxy, and flower development.

#### Closely Related Genes from Other Species

G1792 shows sequence similarity, outside the conserved AP2 domain, with a portion of a predicted protein from tomato, represented by EST sequence AI776626 (AI776626 EST257726 tomato resistant, Cornell Lycopersicon esculentum cDNA clone cLER19A14, mRNA sequence).

G1756: G1756 (SEQ ID NO:175) was studied using transgenic plants in which the gene was expressed under the control of the 35S promoter. Overexpression of G1756 caused alterations in plant growth and development, reducing overall plant size and fertility. In addition, 35S::G1756 overexpressing lines show more disease symptoms following inoculation with a low dose of the fungal pathogen *Botrytis cinerea* compared to the wild-type controls. G1756 was ubiquitously expressed and transcript levels were altered by a variety of environmental or physiological conditions; G1756 expression can be induced by auxin, cold, and *Fusarium*.

#### Closely Related Genes from Other Species

G1756 shows some sequence similarity with known genes from other plant species within the conserved WRKY domain.

G481: Northern blot data from five different tissue samples indicates that G481 (SEQ ID NO: 113) was primarily expressed in flower and/or silique, and root

tissue. G481 was analyzed through its ectopic overexpression in plants. G481 overexpressors were more tolerant to high sucrose in a germination assay. The phenotype of G481 was mild; however, there was a consistent difference in the hypocotyl and root elongation in the overexpressor plants compared to wild-type controls. Sucrose-sensing has been implicated in the regulation of source-sink relationships in plants. Consistent with the sugar sensing phenotype of the G481 overexpressors were the results from the biochemical analysis of G481 overexpressor plants suggesting that line 14 had higher amounts of seed oils and lower amounts of seed protein. This suggested that G481 was involved in the allocation of storage compounds to the seed. G481 overexpressor line 8 was darker green in the T2 generation which could mean a higher photosynthetic rate consistent with the possible role of G481 in sugar sensing.

#### Closely Related Genes from Other Species

There are several sequences from higher plants that show significant homology to G481 including, X59714 from corn, and two ESTs from tomato, AI486503 and AI782351.

G2133: G2133 (SEQ ID NO: 141) was studied using transgenic plants in which the gene was expressed under the control of the 35S promoter. Overexpression of G2133 caused a variety of alterations in plant growth and development: delayed flowering, altered inflorescence architecture, and a decrease in overall size and fertility. At early stages, 35S::G2133 transformants were markedly smaller than controls and displayed curled, dark-green leaves. Most of these plants remained in a vegetative phase of development substantially longer than controls, and produced an increased number of leaves before bolting. In the most severely affected plants, bolting occurred more than a month later than in wild type (24-hour light). In addition, the plants displayed a reduction in apical dominance and formed large numbers of shoots simultaneously, from the axils of rosette leaves. These inflorescence stems had short internodes, and carried increased numbers of cauline leaf nodes, giving them a very leafy appearance. The fertility of 35S::G2133 plants was generally very low. In addition, G2133 overexpressing lines were more resistant to the herbicide glyphosate. In a repeat experiment, lines 4 and 5 were more tolerant

while line 2 was wild-type. G2133 expression was detected in a variety of flower, leaf, embryo, and silique samples. Its expression was altered under various conditions, including auxin treatment, osmotic stress, and Fusarium infection. G2133 can be used for the generation of glyphosate resistant plants, and to increase plant resistance to oxidative stress.

#### Closely Related Genes from Other Species

G2133 shows some sequence similarity with known genes from other plant species within the conserved AP2/EREBP domain.

G2517: G2517 (SEQ ID NO: 143) was studied using transgenic plants in which the gene was expressed under the control of the 35S promoter. Overexpression of G2517 caused alterations in plant growth and development: size variation was apparent in the 35S::G2517 T1 generation, with at least half the lines being very small. Additionally, 4/12 T1 plants formed flower buds marginally earlier than wild type. Three T1 lines (#8,11,12) were examined in the T2 generation, and all three T2 populations were slightly smaller than controls. In the physiological analysis of the T2 populations, G2517 overexpressing lines were more resistant to the herbicide glyphosate. G2517 can be used for the generation of glyphosate resistant plants, and to increase plant resistance to oxidative stress.

#### Closely Related Genes from Other Species

G2517 shows some sequence similarity with known genes from other plant species within the conserved WRKY domain.

G2140: The complete sequence of G2140 (SEQ ID NO: 81) was determined. G2140 was expressed throughout the plant. It showed repression by salicylic acid and Erysiphe infection. Overexpressing G2140 in Arabidopsis resulted in seedlings that were more tolerant to osmotic stress conditions. In germination assays where seedlings were exposed to high concentrations of sucrose or NaCl, all three lines tested showed better cotyledon expansion and seedling vigor. Additionally, G2140 overexpressing plants showed insensitivity to ABA in a germination assay. In

general, G2140 overexpressing plants were small and sickly with shriveled leaves when grown in Petri plates. The combination of ABA insensitivity and resistance to osmotic stress at germination had also been observed for other genes, for example, G1820 (SEQ ID NO:13) and G926 (SEQ ID NO:111). Significantly, the ABA resistance was detected in a germination assay. ABA is involved in maintaining seed dormancy, and it is possible that ABA insensitivity at the germination stage promotes germination despite unfavorable conditions.

When grown in soil, G2140 overexpressing plants displayed marked changes in Arabidopsis leaf and root morphology. All twenty of the 35S::G2140 primary transformants displayed, to various extents, leaves with upcurled margins. In the most severe cases, the leaves became highly contorted and the plants were slightly small and grew more slowly than controls. Three T1 lines (#12, 15 and 16) that showed substantial levels of G2140 overexpression (determined by RT-PCR) were chosen for further study. The T2 seedlings from each of these lines exhibited stunted roots compared with controls. Seedlings from two of the lines (#15,16) also showed upcurled cotyledons. At later stages, however, T2-16 plants appeared wild type. Plants from the T2-12 and T2-15 populations were rather varied in size and showed hints of leaf curling later in development. However, this effect was less severe than that seen in the T1 lines. To verify the leaf-curling phenotype, two further T2 populations (#3,18) were morphologically examined; seedlings from T2-3 were extremely tiny with thickened hypocotyls and short stunted roots. Such plants were too small for transfer to soil. However, T2-18 plants showed slightly contorted cotyledons and formed severely upcurled leaves, confirming the effects seen in the T1 generation.

G2140 is useful for creating plants that germinate better under conditions of high salt. Evaporation from the soil surface causes upward water movement and salt accumulation in the upper soil layer where the seeds are placed. Thus, germination normally takes place at a salt concentration much higher than the mean salt concentration in the whole soil profile. Increased salt tolerance during the germination stage of a crop plant will impact survivability and yield. In addition, G2140 can be used to alter a plant's response to water deficit conditions and,

therefore, can be used to engineer plants with enhanced tolerance to freezing.

#### Closely Related Genes from Other Species

G2140 proteins show extensive sequence similarity with a tomato ovary cDNA, TAMU Lycopersicon esculentum (A1488313) and a Glycine max cDNA clone (BE020519).

G1946: G1946 (SEQ ID NO:49) was studied using transgenic plants in which the gene was expressed under the control of the 35S promoter. Overexpression of G1946 resulted in accelerated flowering, with 35S::G1946 transformants producing flower buds up to a week earlier than wild-type controls (24-hour light conditions). These effects were seen in 12/20 primary transformants and in two independent plantings of each of the three T2 lines. Unlike many early flowering Arabidopsis transgenic lines, which are dwarfed, 35S::G1946 transformants often reached full-size at maturity, and produced large quantities of seeds, although the plants were slightly pale in coloration and had slightly flat leaves compared to wild-type. In addition, 35S::G1946 plants showed an altered response to phosphate deprivation. Seedlings of G1946 overexpressors showed more secondary root growth on phosphate-free media, when compared to wild-type control. In a repeat experiment, all three lines showed the phenotype. Overexpression of G1946 in Arabidopsis also resulted in an increase in seed glucosinolate M39501 in T2 lines 1 and 3. An increase in seed oil and a decrease in seed protein was also observed in these two lines. G1946 was ubiquitously expressed, and did not appear to be significantly induced or repressed by any of the biotic and abiotic stress conditions tested, with the exception of cold, which repressed G1946 expression. G1946 can be used to modify flowering time, as well as to improve the plant's performance in conditions of limited phosphate, and to alter seed oil, protein, and glucosinolate composition.

#### Closely Related Genes from Other Species

A comparison of the amino acid sequence of G1946 with sequences available from GenBank showed strong similarity with plant HSFs of several species

(*Lycopersicon peruvianum*, *Medicago truncatula*, *Lycopersicon esculentum*, *Glycine max*, *Solanum tuberosum*, *Oryza sativa* and *Hordeum vulgare* subsp.

G1852: G1852 (SEQ ID NO:71) was analyzed through its ectopic overexpression in plants. Analysis of the endogenous level of G1852 transcripts by RT-PCR revealed expression in all tissues tested. G1852 expression was induced in response to ABA, heat and drought treatment. 35S::G1852 overexpressor plants were more tolerant to osmotic stress in a root growth assay on PEG (polyethylene glycol)-containing media compared with wild-type controls. Seedlings were slightly larger and have more root growth. G1852 can be used to alter a plant's response to water deficit conditions and therefore, be used to engineer plants with enhanced tolerance to drought, salt stress, and freezing.

#### Closely Related Genes from Other Species

A comparison of the amino acid sequence of G1852 with entries available from GenBank shows strong similarity with plant ankyrins of several species (*Malus domestica*, *Solanum tuberosum*, *Oryza sativa*, *Gossypium arboreum*, *Medicago truncatula*, *Glycine max*, *Lycopersicon esculentum*, *Pinus taeda*, *Lotus japonicus* and *Gossypium hirsutum*).

G325: G325 (SEQ ID NO:95) was analyzed using transgenic plants in which G325 was expressed under the control of the 35S promoter. G325 overexpressing plants showed more tolerance to osmotic stress in a germination assay in three separate experiments. They showed more seedling vigor than wild-type control when germinated on plates containing high salt and high sucrose. G325 was expressed at high levels in flowers and cauline leaves, and at lower levels in shoots, rosette leaves, and seedlings. G325 was induced by auxin, cold and heat stress. Expression of G325 was reduced in response to *Fusarium* infection or salicylic acid treatment. G325 can be useful for enhancing seed germination under high salt conditions or other conditions of osmotic stress. Evaporation from the soil surface causes upward water movement and salt accumulation in the upper soil layer where the seeds are placed. Thus, germination normally takes place at a salt concentration much higher than the mean salt concentration of in the whole soil profile. Increased salt tolerance during the germination stage of a crop plant will impact survivability and yield. G325 can



also be used to engineer plants with enhanced tolerance to drought, freezing at later stages during growth and development.

#### Closely Related Genes from Other Species

G325 showed homology to non-Arabidopsis proteins within the conserved domain.

G2583: G2583 (SEQ ID NO: 17) was studied using transgenic plants in which the gene was expressed under the control of the 35S promoter. 35S::G2583 plants exhibited extremely glossy leaves. At early stages, 35S::G2583 seedlings appeared normal, but by about two weeks after sowing, the plants exhibited very striking shiny leaves, which were apparent until very late in development. Many lines displayed a variety of other effects such as a reduction in overall size, narrow curled leaves, or various non-specific floral abnormalities, which reduced fertility. These effects on leaf appearance were observed in 18/20 primary transformants, and in all the plants from 4/6 of the T2 lines (#2,4,9 and 15) examined. The glossy nature of the leaves from 35S::G2583 plants can be a consequence of changes in epicuticular wax content or composition. G2583 belongs to a small clade within the large AP2/EREBP Arabidopsis family that also contains G975 (SEQ ID NO: 19), G1387 (SEQ ID NO: 21), and G977 (SEQ ID NO: 23). Overexpression of G975 caused a substantial increase in leaf wax components, as well as morphological phenotypes resembling those observed in 35S::G2583 plants. G2583 was ubiquitously expressed, at higher levels in root, flower, embryo, and silique tissues. G2583 can be used to modify plant appearance (shiny leaves). In addition, it can be used to manipulate wax composition, amount, or distribution, which in turn can modify plant tolerance to drought and/or low humidity or resistance to insects.

#### Closely Related Genes from Other Species

G2583 showed some sequence similarity with known genes from other plant species within the conserved AP2/EREBP domain.

G1322: G1322 (SEQ ID NO: 21) was analyzed using transgenic plants. The gene was expressed under the control of the 35S promoter. 35S transgenic plants were wild-type in phenotype with respect to the biochemical analyses performed. Overexpression of G1322 produced changes in overall plant size and leaf development. At all stages, 35S::G1322 plants were distinctly smaller than controls and developed curled dark-green leaves. Following the switch to flowering, the plants formed relatively thin inflorescence stems and had a rather poor seed yield. In addition, overexpression of G1322 resulted in plants with an altered etiolation response as well as enhanced tolerance to germination under chilling conditions. When germinated in the dark, G1322 overexpressing transgenic plant lines had open, slightly green cotyledons. Under chilling conditions, all three transgenic lines displayed a similar germination response, seedlings were slightly larger and had longer roots. In addition, an increase in the leaf glucosinolate M39480 was observed in all three T2 lines. According to RT-PCR analysis, G1322 was expressed primarily in flower tissue. The utilities of G1322 include altering a plant's chilling sensitivity and altering a plant's light response. The germination of many crops is very sensitive to cold temperatures. A gene that will enhance germination and seedling vigor in the cold has tremendous utility in allowing seeds to be planted earlier in the season with a higher survival rate. G1322 can also be useful for altering leaf glucosinolate composition. Increases or decreases in specific glucosinolates or total glucosinolate content are desirable depending upon the particular application. Modification of glucosinolate composition or quantity can therefore afford increased protection from predators. Furthermore, in edible crops, tissue specific promoters can be used to ensure that these compounds accumulate specifically in tissues, such as the epidermis, which are not taken for consumption.

#### Closely Related Genes from Other Species

G1322 shows some sequence similarity with known genes from other plant species within the conserved Myb domain.

G303: The complete sequence of G303 (SEQ ID NO: 23) was determined. G303 was detected at very low levels in roots and rosette leaves. G303 was analyzed using transgenic plants in which G303 was expressed under the control of the 35S

promoter. G303 overexpressing plants showed more tolerance to osmotic stress in a germination assay in three separate experiments. They showed more germination than wild-type control when germinated on plates containing high salt and high sucrose. G303 are useful for enhancing seed germination under high salt conditions or other conditions of osmotic stress. Evaporation from the soil surface causes upward water movement and salt accumulation in the upper soil layer where the seeds are placed. Thus, germination normally takes place at a salt concentration much higher than the mean salt concentration in the whole soil profile. Increased salt tolerance during the germination stage of a crop plant will impact survivability and yield. G303 can also be used to engineer plants with enhanced tolerance to drought, salt stress, and freezing.

#### Closely Related Genes from Other Species

G303 shows some sequence similarity with known genes from other plant species within the conserved basic HLH domain.

G1927:G1927 (SEQ ID NO: 239) was analyzed using transgenic plants in which the gene was expressed under the control of the 35S promoter. Overexpression of G1927 in Arabidopsis resulted in plants that had an altered response to pathogen. Plants overexpressing G1927 showed fewer disease symptoms following infection with the fungal pathogen *Sclerotinia sclerotiorum* compared with control plants. The experiment was repeated on individual lines, and all three lines showed the enhanced pathogen tolerance phenotype. G1927 expression appeared to be ubiquitous according to RT-PCR analysis. G1927 can be used to manipulate the defense response in order to generate pathogen-resistant plants.

#### Closely Related Genes from Other Species

G1927 showed extensive sequence similarity to a NAC protein from tomato (BG350410).

**Example VIII: Identification of Homologous Sequences**

Homologous sequences from *Arabidopsis* and plant species of *Arabidopsis* were identified using database sequence search tools, such as the Basic Local Alignment Search Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; and Altschul et al. (1997) Nucl. Acid Res. 25: 3389-3402). The tblastx sequence analysis programs were employed using the BLOSUM-62 scoring matrix (Henikoff, S. and Henikoff, J. G. (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919).

Identified non-*Arabidopsis* sequences homologous to the *Arabidopsis* sequences are provided in Table 5. The percent sequence identity among these sequences can be as low as 47%, or even lower sequence identity. The entire NCBI GenBank database was filtered for sequences from all plants except *Arabidopsis thaliana* by selecting all entries in the NCBI GenBank database associated with NCBI taxonomic ID 33090 (Viridiplantae; all plants) and excluding entries associated with taxonomic ID 3701 (*Arabidopsis thaliana*). These sequences are compared to sequences representing genes of SEQ IDs NOs:2 - 2N, where N = 2-123, using the Washington University TBLASTX algorithm (version 2.0a19MP) at the default settings using gapped alignments with the filter "off". For each gene of SEQ IDs NOs:2 - 2N, where N = 2-123, individual comparisons were ordered by probability score (P-value), where the score reflects the probability that a particular alignment occurred by chance. For example, a score of  $3.6e-40$  is  $3.6 \times 10^{-40}$ . In addition to P-values, comparisons were also scored by percentage identity. Percentage identity reflects the degree to which two segments of DNA or protein are identical over a particular length. Examples of sequences so identified are presented in Table 5. Homologous or orthologous sequences are readily identified and available in GenBank by Accession number (Table 5; Test sequence ID). The identified homologous polynucleotide and polypeptide sequences and homologues of the *Arabidopsis* polynucleotides and polypeptides may be orthologs of the *Arabidopsis* polynucleotides and polypeptides.

**Example IX Introduction of polynucleotides into dicots**

SEQ ID NOs:1-(2N - 1), wherein N = 2-123, paralogous, orthologous, and homologous sequences recombined into pMEN20 or pMEN65 expression vectors are transformed into a plant for the purpose of modifying plant traits. The cloning vector may be introduced into a variety of cereal plants by means well-known in the art such

as, for example, direct DNA transfer or *Agrobacterium tumefaciens*-transformation. It is now routine to produce transgenic plants using (see Weissbach and Weissbach, (1989) *supra*; Gelvin et al., (1990) *supra*; Herrera-Estrella et al. (1983) *supra*; Bevan (1984) *supra*; and Klee (1985) *supra*). Methods for analysis of traits are routine in the art and examples are disclosed above.

#### **Example X Transformation of Cereal Plants with an Expression Vector**

Cereal plants such as corn, wheat, rice, sorghum or barley, may also be transformed with the present polynucleotide sequences in pMEN20 or pMEN65 expression vectors for the purpose of modifying plant traits. For example, pMEN020 may be modified to replace the NptII coding region with the BAR gene of *Streptomyces hygroscopicus* that confers resistance to phosphinothricin. The KpnI and BglII sites of the Bar gene are removed by site-directed mutagenesis with silent codon changes.

The cloning vector may be introduced into a variety of cereal plants by means well-known in the art such as, for example, direct DNA transfer or *Agrobacterium tumefaciens*-mediated transformation. It is now routine to produce transgenic plants of most cereal crops (Vasil, I., Plant Molec. Biol. 25: 925-937 (1994)) such as corn, wheat, rice, sorghum (Cassas, A. et al., Proc. Natl. Acad Sci USA 90: 11212-11216 (1993) and barley (Wan, Y. and Lemeaux, P. Plant Physiol. 104:37-48 (1994). DNA transfer methods such as the microprojectile can be used for corn (Fromm. et al. Bio/Technology 8: 833-839 (1990); Gordon-Kamm et al. Plant Cell 2: 603-618 (1990); Ishida, Y., Nature Biotechnology 14:745-750 (1990)), wheat (Vasil, et al. Bio/Technology 10:667-674 (1992) ; Vasil et al., Bio/Technology 11:1553-1558 (1993); Weeks et al., Plant Physiol. 102:1077-1084 (1993)), rice (Christou Bio/Technology 9:957-962 (1991); Hiei et al. Plant J. 6:271-282 (1994); Aldemita and Hodges, Planta 199:612-617; Hiei et al., Plant Mol Biol. 35:205-18 (1997)). For most cereal plants, embryogenic cells derived from immature scutellum tissues are the preferred cellular targets for transformation (Hiei et al., Plant Mol Biol. 35:205-18 (1997); Vasil, Plant Molec. Biol. 25: 925-937 (1994)).

Vectors according to the present invention may be transformed into corn embryogenic cells derived from immature scutellar tissue by using microprojectile

bombardment, with the A188XB73 genotype as the preferred genotype. al., *Bio/Technology* 8: 833-839 (1990); Gordon-Kamm et al., *Plant* (1990)). After microprojectile bombardment the tissues are selected on phosphinothricin to identify the transgenic embryogenic cells (Gordon-Kamm et al., *Plant Cell* 2: 603-618 (1990)). Transgenic plants are regenerated by standard corn regeneration techniques (Fromm, et al., *Bio/Technology* 8: 833-839 (1990); Gordon-Kamm et al., *Plant Cell* 2: 603-618 (1990)).

The plasmids prepared as described above can also be used to produce transgenic wheat and rice plants (Christou, *Bio/Technology* 9:957-962 (1991); Hiei et al., *Plant J.* 6:271-282 (1994); Aldemita and Hodges, *Planta* 199:612-617 (1996); Hiei et al., *Plant Mol Biol.* 35:205-18 (1997)) that coordinately express genes of interest by following standard transformation protocols known to those skilled in the art for rice and wheat Vasil, et al. *Bio/Technology* 10:667-674 (1992) ; Vasil et al., *Bio/Technology* 11:1553-1558 (1993); Weeks et al., *Plant Physiol.* 102:1077-1084 (1993)), where the bar gene is used as the selectable marker.

All references, publications, patent documents, web pages, and other documents cited or mentioned herein are hereby incorporated by reference in their entirety for all purposes. Although the invention has been described with reference to specific embodiments and examples, it should be understood that one of ordinary skill can make various modifications without departing from the spirit of the invention. The scope of the invention is not limited to the specific embodiments and examples provided.

We claim:

1. A transgenic plant comprising a recombinant polynucleotide having a nucleotide sequence selected from the group consisting of:
  - (a) a nucleotide sequence encoding a polypeptide comprising a sequence selected from those of, SEQ ID NOs: 26, 46, 176, 114, 142, 144, 82, 50, 72, 96, 18, 22, 24, 240, or a complementary nucleotide sequence thereof;
  - (b) a nucleotide sequence of SEQ ID NOs: 25, 45, 175, 113, 141, 143, 81, 49, 71, 95, 17, 21, 23, 239, or a complementary nucleotide sequence thereof; and
  - (c) a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence of one or more polynucleotides of: (a) or (b).
2. The transgenic plant of claim 1 wherein the transgenic plant possesses an altered trait as compared to another plant, or the transgenic plant exhibits an altered phenotype as compared to another plant, or the transgenic plant expresses an altered level of one or more genes associated with a plant trait as compared to another plant, wherein the other plant does not comprise the recombinant polynucleotide.
3. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is an enhanced tolerance to environmental conditions.
4. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is an improved tolerance to microbial, fungal or viral diseases.
5. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is an improved tolerance to pest infestations.
6. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is a decreased herbicide sensitivity.

7. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is an improved tolerance or enhanced ability to take up heavy metals.
8. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is improved growth under poor photoconditions.
9. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is improved nutrient uptake.
10. The transgenic plant of claim 1 wherein the plant possesses an altered trait as compared to another plant wherein the trait is reduced hormone sensitivity.
11. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:26.
12. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:46.
13. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:176.
14. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:114.
15. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:142.
16. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:144.
17. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:82.



18. The transgenic plant of claim 1, wherein the recombinant poly comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:50.
19. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:72.
20. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:96.
21. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:18.
22. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:22.
23. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:24.
24. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:240.
25. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:25.
26. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:45.
27. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:175.
28. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:113.

29. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:141.
30. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:143.
31. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:81.
32. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:49.
33. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:71.
34. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:95.
35. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:17.
36. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:21.
37. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:23.
38. The transgenic plant of claim 1, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:239.
39. The transgenic plant of claim 1, further comprising a constitutive, inducible, or tissue-specific promoter operably linked to said nucleotide sequence.

40. The transgenic plant of claim 1, wherein the plant is selected from consisting of: soybean, wheat, corn, potato, cotton, rice, oilseed rape, alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honeydew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, pumpkin, spinach, squash, sweet corn, tobacco, tomato, watermelon, mint and other labiates, rosaceous fruits, and vegetable brassicas.
41. The transgenic plant of claim 1 wherein the encoded polypeptide is expressed and regulates transcription of a gene.
42. A method of using the transgenic plant of claim 1 to grow a progeny plant from a parent plant, the method comprising crossing the transgenic plant with another plant, selecting seed, and growing the progeny plant from the seed.
43. An isolated or recombinant polynucleotide comprising a nucleotide sequence selected from the group consisting of:
- (a) a nucleotide sequence encoding a polypeptide comprising a sequence selected from SEQ ID NOs:26, 82, 24, 240, or a complementary nucleotide sequence thereof;
  - (b) a nucleotide sequence of SEQ ID NOs:25, 81, 23, 239, or a complementary nucleotide sequence thereof; and
  - (c) a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence of one or more of: (a) or (b).
44. The isolated or recombinant polynucleotide of claim 43, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:26.
45. The isolated or recombinant polynucleotide of claim 43, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:82.

46. The isolated or recombinant polynucleotide of claim 43, where recombinant polynucleotide comprises a nucleotide sequence encoded of SEQ ID NO:24.
47. The isolated or recombinant polynucleotide of claim 43, wherein the recombinant polynucleotide comprises a nucleotide sequence encoding a polypeptide of SEQ ID NO:240.
48. The isolated or recombinant polynucleotide of claim 43, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:26.
49. The isolated or recombinant polynucleotide of claim 43, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:81.
50. The isolated or recombinant polynucleotide of claim 43, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:23.
51. The isolated or recombinant polynucleotide of claim 43, wherein the recombinant polynucleotide comprises a nucleotide sequence of SEQ ID NO:239.
52. The isolated or recombinant polynucleotide of claim 43, further comprising a constitutive, inducible, or tissue-specific promoter operably linked to the nucleotide sequence.
53. The isolated or recombinant polynucleotide of claim 43 wherein the encoded polypeptide is expressed and regulates transcription of a gene.
54. A vector comprising the isolated or recombinant polynucleotide of claim 43.
55. A host cell comprising the vector of claim 54.
56. A method of using the isolated or recombinant polynucleotide of claim 43 for producing a modified plant having a modified trait, the method comprising selecting a

polynucleotide that encodes a polypeptide, inserting the polynucleotide expression vector, introducing the vector into a plant or a cell of a plant, overexpress the polypeptide, thereby producing a modified plant, and selecting for a modified trait.

57. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is an enhanced tolerance to environmental conditions.

58. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is an improved tolerance to microbial, fungal or viral diseases.

59. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is an improved tolerance to pest infestations.

60. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is a decreased herbicide sensitivity.

61. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is an improved tolerance of heavy metals or enhanced ability to take up heavy metals.

62. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is improved growth under poor photoconditions.

63. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is improved nutrient uptake.

64. The method of claim 54 wherein the modified plant possesses a modified trait as compared to another plant wherein the trait is reduced hormone sensitivity.

65. A modified plant produced by the method of claim 54.
66. A method of using the plant of claim 63 to grow a progeny plant from a parent plant, the method comprising crossing the transgenic plant with another plant, selecting seed, and growing the progeny plant from the seed.

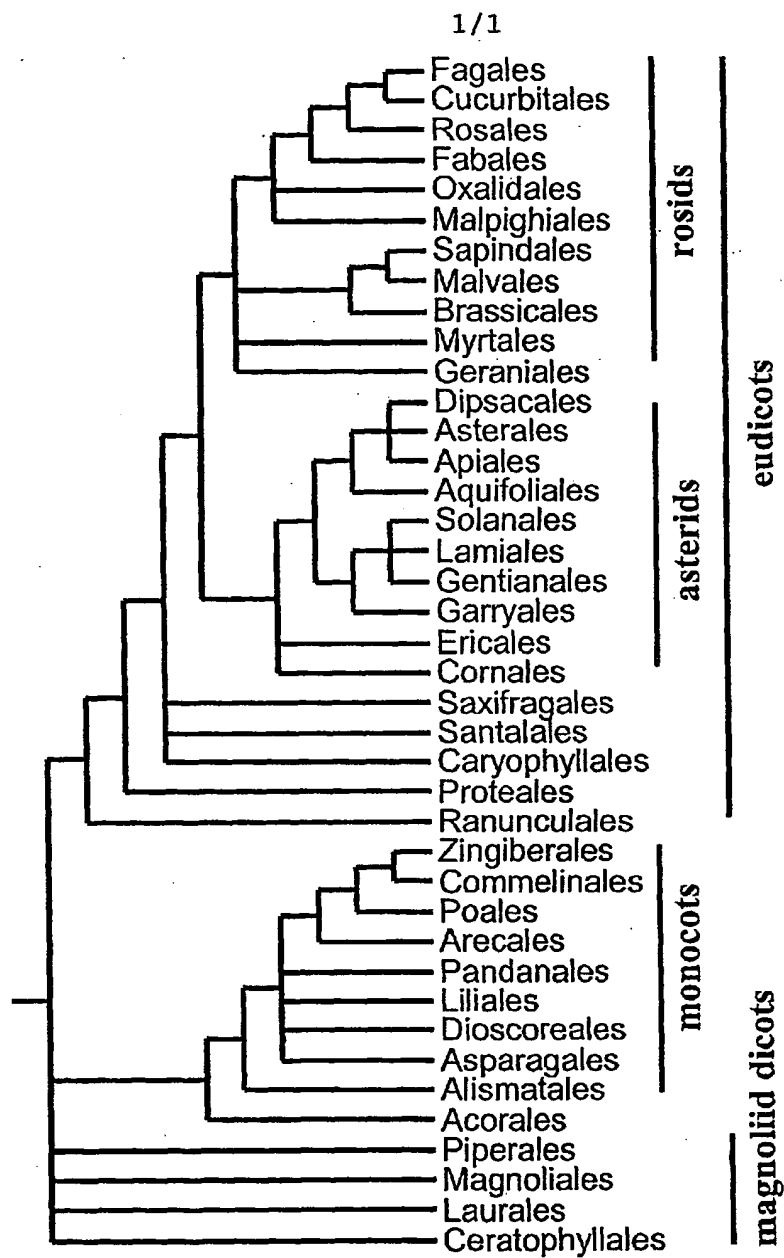


Figure 1

## SEQUENCE LISTING

&lt;110&gt; Mendel Biotechnology, Inc.

Reuber, T. Lynne

Riechmann, Jose Luis

Heard, Jacqueline E.

Jiang, Cai-Zhong

Adam, Luc J.

Dubell, Arnold T.

Ratcliffe, Oliver

Pineda, Omaira

Yu, Guo-Liang

Broun, Pierre E.

<120> STRESS-RELATED POLYNUCLEOTIDES AND  
POLYPEPTIDES IN PLANTS

&lt;130&gt; 514442002040

&lt;150&gt; 60/310,847

&lt;151&gt; 2001-08-09

&lt;150&gt; 60/336,049

&lt;151&gt; 2001-11-19

&lt;150&gt; 60/338,692

&lt;151&gt; 2001-12-11

&lt;150&gt; 10/171,468

&lt;151&gt; 2002-06-14

&lt;160&gt; 246

&lt;170&gt; FastSEQ for Windows Version 4.0

&lt;210&gt; 1

&lt;211&gt; 1011

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (213)...(833)

&lt;400&gt; 1

```

aaagttattg atagtttctg ttacttatta atttttaagg ttatgtgtat tattaccaat 60
tggaggacta tatagtcgca agtctcaacc ctataaaaga aaacattcgt cgatcatctt 120
ccgcctcga gtatctctct ctctctctct cttctctggt ttctttattg attgcataga 180
caaaaataca cacatacaca acagaaagaa ag atg gag acg acg atg aag aag 233
                               Met Glu Thr Thr Met Lys Lys
                               1           5

```

```

aaa ggg aga gtg aaa gcg aca ata acg tca cag aaa gaa gaa gaa gga 281
Lys Gly Arg Val Lys Ala Thr Ile Thr Ser Gln Lys Glu Glu Glu Gly
      10              15              20

```

```

aca gtg aga aaa gga cct tgg act atg gaa.gaa gat ttc atc ctc ttt 329

```



```

Thr Val Arg Lys Gly Pro Trp Thr Met Glu Glu Asp Phe Ile Leu Phe
  25                      30                      35

aat tac atc ctt aat cat ggt gaa ggt ctt tgg aac tct gtc gcc aaa 377
Asn Tyr Ile Leu Asn His Gly Glu Gly Leu Trp Asn Ser Val Ala Lys
  40                      45                      50                      55

gcc tct ggt cta aaa cgt act gga aaa agt tgt cgg ctc cgg tgg ctg 425
Ala Ser Gly Leu Lys Arg Thr Gly Lys Ser Cys Arg Leu Arg Trp Leu
                      60                      65                      70

aac tat ctc cga cca gat gtg cgg cga ggg aac ata acc gaa gaa gaa 473
Asn Tyr Leu Arg Pro Asp Val Arg Arg Gly Asn Ile Thr Glu Glu Glu
                      75                      80                      85

cag ctt ttg atc att cag ctt cat gct aag ctt gga aac agg tgg tcg 521
Gln Leu Leu Ile Ile Gln Leu His Ala Lys Leu Gly Asn Arg Trp Ser
                      90                      95                      100

aag att gcg aag cat ctt ccg gga aga acg gac aac gag ata aag aac 569
Lys Ile Ala Lys His Leu Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn
  105                      110                      115

ttc tgg agg aca aag att cag aga cac atg aaa gtg tca tcg gaa aat 617
Phe Trp Arg Thr Lys Ile Gln Arg His Met Lys Val Ser Ser Glu Asn
  120                      125                      130                      135

atg atg aat cat caa cat cat tgt tcg gga aac tca cag agc tcg ggg 665
Met Met Asn His Gln His His Cys Ser Gly Asn Ser Gln Ser Ser Gly
                      140                      145                      150

atg acg acg caa ggc agc tcc ggc aaa gcc ata gac acg gct gag agc 713
Met Thr Thr Gln Gly Ser Ser Gly Lys Ala Ile Asp Thr Ala Glu Ser
                      155                      160                      165

ttc tct cag gcg aag acg acg acg ttt aat gtg gtg gaa caa cag tca 761
Phe Ser Gln Ala Lys Thr Thr Thr Phe Asn Val Val Glu Gln Gln Ser
                      170                      175                      180

aac gag aat tac tgg aac gtt gaa gat ctg tgg ccc gtc cac ttg ctt 809
Asn Glu Asn Tyr Trp Asn Val Glu Asp Leu Trp Pro Val His Leu Leu
                      185                      190                      195

aat ggt gac cac cat gtg att taa gatatatata tagacctcct atacatttat 863
Asn Gly Asp His His Val Ile *
  200                      205

atgccccagc tgggtttttt tgtatggtac gttatttggg ttttctattg ctgaaatgtc 923
gttgcatatta atttacatac gaaaagtga ttaaatcatt aaatcttcaa tacatatgga 983
ggtggtgttt gagtaaaaaa aaaaaaaa
                      1011

<210> 2
<211> 206
<212> PRT
<213> Arabidopsis thaliana

<220>

```

<221> DOMAIN  
 <222> (26)...(130)  
 <223> Conserved domain

<400> 2

```

Met Glu Thr Thr Met Lys Lys Lys Gly Arg Val Lys Ala Thr Ile Thr
 1           5           10           15
Ser Gln Lys Glu Glu Gly Thr Val Arg Lys Gly Pro Trp Thr Met
           20           25           30
Glu Glu Asp Phe Ile Leu Phe Asn Tyr Ile Leu Asn His Gly Glu Gly
           35           40           45
Leu Trp Asn Ser Val Ala Lys Ala Ser Gly Leu Lys Arg Thr Gly Lys
           50           55           60
Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp Val Arg Arg
65           70           75           80
Gly Asn Ile Thr Glu Glu Gln Leu Leu Ile Ile Gln Leu His Ala
           85           90           95
Lys Leu Gly Asn Arg Trp Ser Lys Ile Ala Lys His Leu Pro Gly Arg
           100          105          110
Thr Asp Asn Glu Ile Lys Asn Phe Trp Arg Thr Lys Ile Gln Arg His
           115          120          125
Met Lys Val Ser Ser Glu Asn Met Met Asn His Gln His His Cys Ser
           130          135          140
Gly Asn Ser Gln Ser Ser Gly Met Thr Thr Gln Gly Ser Ser Gly Lys
145          150          155          160
Ala Ile Asp Thr Ala Glu Ser Phe Ser Gln Ala Lys Thr Thr Thr Phe
           165          170          175
Asn Val Val Glu Gln Gln Ser Asn Glu Asn Tyr Trp Asn Val Glu Asp
           180          185          190
Leu Trp Pro Val His Leu Leu Asn Gly Asp His His Val Ile
           195          200          205

```

<210> 3  
 <211> 1019  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (41)...(988)

<400> 3

```

cctctcttca ttttttaact ccctctctct ctctctctct atg gag aga cga acg 55
                                   Met Glu Arg Arg Thr
                                   1           5

aga cga gtg aag ttc aca gag aat cgt acg gtc aca aac gta gca gct 103
Arg Arg Val Lys Phe Thr Glu Asn Arg Thr Val Thr Asn Val Ala Ala
           10           15           20

aca cca tct aac ggg tct ccg aga ctg gtc cgt atc act gtt act gat 151
Thr Pro Ser Asn Gly Ser Pro Arg Leu Val Arg Ile Thr Val Thr Asp
           25           30           35

cct ttc gct act gac tcg tct agc gac gac gac aac aac aac gtc 199
Pro Phe Ala Thr Asp Ser Ser Ser Asp Asp Asp Asn Asn Asn Val

```

40	45	50	
acg gtt gtt cca aga gtg aaa cga tac gtg aag gag att aga ttc tgc			247
Thr Val Val Pro Arg Val Lys Arg Tyr Val Lys Glu Ile Arg Phe Cys			
55	60	65	
caa ggt gaa tct tct tcc tcc acc gcg gcg agg aaa ggt aag cac aag			295
Gln Gly Glu Ser Ser Ser Ser Thr Ala Ala Arg Lys Gly Lys His Lys			
70	75	80	85
gag gag gaa agc gta gtg gtt gaa gat gac gtg tcg acg tcg gtg aag			343
Glu Glu Glu Ser Val Val Val Glu Asp Asp Val Ser Thr Ser Val Lys			
90	95	100	
cct aaa aag tac aga ggc gtg aga cag aga cct tgg gga aaa ttc gcg			391
Pro Lys Lys Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Phe Ala			
105	110	115	
gcg gag att aga gat ccg tcg agc cgt act cgg att tgg ctt ggg act			439
Ala Glu Ile Arg Asp Pro Ser Ser Arg Thr Arg Ile Trp Leu Gly Thr			
120	125	130	
ttt gtc acg gcg gag gaa gct gct ata gcg tac gat aga gcc gcg att			487
Phe Val Thr Ala Glu Glu Ala Ala Ile Ala Tyr Asp Arg Ala Ala Ile			
135	140	145	
cat ctc aaa gga cct aaa gcg ctc acg aat ttc cta act ccg ccg acg			535
His Leu Lys Gly Pro Lys Ala Leu Thr Asn Phe Leu Thr Pro Pro Thr			
150	155	160	165
cca acg ccg gtt atc gat ctc caa acg gtt tcc gcc tgc gat tac ggt			583
Pro Thr Pro Val Ile Asp Leu Gln Thr Val Ser Ala Cys Asp Tyr Gly			
170	175	180	
aga gat tct cgg cag agc ctt cat tca ccg acc tct gtt cta aga ttc			631
Arg Asp Ser Arg Gln Ser Leu His Ser Pro Thr Ser Val Leu Arg Phe			
185	190	195	
aac gtc aac gag gaa aca gag cat gag att gaa gcg atc gag cta tct			679
Asn Val Asn Glu Glu Thr Glu His Glu Ile Glu Ala Ile Glu Leu Ser			
200	205	210	
ccg gag aga aag tcg acg gtt ata aaa gaa gaa gaa gaa tcg tcg gcg			727
Pro Glu Arg Lys Ser Thr Val Ile Lys Glu Glu Glu Glu Ser Ser Ala			
215	220	225	
ggt ttg gtg ttc ccg gat ccg tat ctg tta ccg gat tta tct ctc gcc			775
Gly Leu Val Phe Pro Asp Pro Tyr Leu Leu Pro Asp Leu Ser Leu Ala			
230	235	240	245
ggc gaa tgt ttt tgg gat acc gaa att gcc cct gac ctt ttg ttt ctc			823
Gly Glu Cys Phe Trp Asp Thr Glu Ile Ala Pro Asp Leu Leu Phe Leu			
250	255	260	
gat gaa gaa acc aaa atc caa tca acg ttg tta cca aac aca gag gtt			871
Asp Glu Glu Thr Lys Ile Gln Ser Thr Leu Leu Pro Asn Thr Glu Val			
265	270	275	

tcg aaa caa gga gaa aac gaa act gaa gat ttc gag ttt ggt ttg att 919  
 Ser Lys Gln Gly Glu Asn Glu Thr Glu Asp Phe Glu Phe Gly Leu Ile  
 280 285 290

gat gat ttc gag tct tct cca tgg gat gtg gat cat ttc ttc gac cat 967  
 Asp Asp Phe Glu Ser Ser Pro Trp Asp Val Asp His Phe Phe Asp His  
 295 300 305

cat cat cac tct ttc gat taa aaatctcttc ttttttgggg aaatttttgt 1018  
 His His His Ser Phe Asp \*  
 310 315

g 1019

<210> 4

<211> 315

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (93)...(160)

<223> Conserved domain

<400> 4

Met Glu Arg Arg Thr Arg Arg Val Lys Phe Thr Glu Asn Arg Thr Val  
 1 5 10 15  
 Thr Asn Val Ala Thr Pro Ser Asn Gly Ser Pro Arg Leu Val Arg  
 20 25 30  
 Ile Thr Val Thr Asp Pro Phe Ala Thr Asp Ser Ser Ser Asp Asp  
 35 40 45  
 Asp Asn Asn Asn Val Thr Val Val Pro Arg Val Lys Arg Tyr Val Lys  
 50 55 60  
 Glu Ile Arg Phe Cys Gln Gly Glu Ser Ser Ser Thr Ala Ala Arg  
 65 70 75 80  
 Lys Gly Lys His Lys Glu Glu Glu Ser Val Val Val Glu Asp Asp Val  
 85 90 95  
 Ser Thr Ser Val Lys Pro Lys Lys Tyr Arg Gly Val Arg Gln Arg Pro  
 100 105 110  
 Trp Gly Lys Phe Ala Ala Glu Ile Arg Asp Pro Ser Ser Arg Thr Arg  
 115 120 125  
 Ile Trp Leu Gly Thr Phe Val Thr Ala Glu Glu Ala Ala Ile Ala Tyr  
 130 135 140  
 Asp Arg Ala Ala Ile His Leu Lys Gly Pro Lys Ala Leu Thr Asn Phe  
 145 150 155 160  
 Leu Thr Pro Pro Thr Pro Thr Pro Val Ile Asp Leu Gln Thr Val Ser  
 165 170 175  
 Ala Cys Asp Tyr Gly Arg Asp Ser Arg Gln Ser Leu His Ser Pro Thr  
 180 185 190  
 Ser Val Leu Arg Phe Asn Val Asn Glu Glu Thr Glu His Glu Ile Glu  
 195 200 205  
 Ala Ile Glu Leu Ser Pro Glu Arg Lys Ser Thr Val Ile Lys Glu Glu  
 210 215 220  
 Glu Glu Ser Ser Ala Gly Leu Val Phe Pro Asp Pro Tyr Leu Leu Pro  
 225 230 235 240  
 Asp Leu Ser Leu Ala Gly Glu Cys Phe Trp Asp Thr Glu Ile Ala Pro

```

                245                250                255
Asp Leu Leu Phe Leu Asp Glu Glu Thr Lys Ile Gln Ser Thr Leu Leu
                260                265                270
Pro Asn Thr Glu Val Ser Lys Gln Gly Glu Asn Glu Thr Glu Asp Phe
                275                280                285
Glu Phe Gly Leu Ile Asp Asp Phe Glu Ser Ser Pro Trp Asp Val Asp
                290                295                300
His Phe Phe Asp His His His His Ser Phe Asp
305                310                315

```

&lt;210&gt; 5

&lt;211&gt; 1547

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (312)...(1310)

&lt;400&gt; 5

```

tcgtgagcgt tgtgtttctc ctcaacattc aaagtcttta gtgaaacctc tcttgtaaga 60
agccaaaaaa ataaagagaa agattcaaag aaggaaagaa attgaggatg actatttcaa 120
gtccaaagag agattttgag tagaccctct tcacaaaaat ccaatcttag agtcttacta 180
gttactatct agcttacata cacagagaca ctataccaaa aatccaatct tattagagta 240
cttactatat agcttacaca tacacacaca cgaagtacta tttcaacgat caagagcgtg 300
tgcgtgagga t atg ggt aga cca cct tgt tgc gag aag att gag gtg aag 350
      Met Gly Arg Pro Pro Cys Cys Glu Lys Ile Glu Val Lys
              1              5              10

aaa gga cca tgg act ccc gaa gaa gac ata atc ttg gtc tct tat atc 398
Lys Gly Pro Trp Thr Pro Glu Glu Asp Ile Ile Leu Val Ser Tyr Ile
      15              20              25

caa caa cac ggc cct gga aat tgg aga tct gtc cct gca aac acc ggt 446
Gln Gln His Gly Pro Gly Asn Trp Arg Ser Val Pro Ala Asn Thr Gly
      30              35              40              45

ttg cta agg tgt agc aag agt tgc aga ctt aga tgg act aat tac ctt 494
Leu Leu Arg Cys Ser Lys Ser Cys Arg Leu Arg Trp Thr Asn Tyr Leu
              50              55              60

cgt ccc ggg atc aaa cga gga aat ttc act caa ccg gaa gag aag atg 542
Arg Pro Gly Ile Lys Arg Gly Asn Phe Thr Gln Pro Glu Glu Lys Met
              65              70              75

atc atc cac ctt caa gct ctt ttg gga aat aga tgg gca gct ata gca 590
Ile Ile His Leu Gln Ala Leu Leu Gly Asn Arg Trp Ala Ala Ile Ala
              80              85              90

tca tat cta cct cag agg acc gac aat gat atc aag aac tac tgg aac 638
Ser Tyr Leu Pro Gln Arg Thr Asp Asn Asp Ile Lys Asn Tyr Trp Asn
              95              100              105

act cat ctt aaa aag aaa cta gtg atg atg aag ttt caa aat ggt atc 686
Thr His Leu Lys Lys Lys Leu Val Met Met Lys Phe Gln Asn Gly Ile
110              115              120              125

```

atc aac gaa aac aaa acc aat ctg gca aca gat att tcg tct tgt aat	734
Ile Asn Glu Asn Lys Thr Asn Leu Ala Thr Asp Ile Ser Ser Cys Asn	
130 135 140	
aat aac aac aat gga tgt aat cac aac aaa agg acc acc aac aaa ggc	782
Asn Asn Asn Asn Gly Cys Asn His Asn Lys Arg Thr Thr Asn Lys Gly	
145 150 155	
caa tgg gag aaa aaa ctt caa aca gac atc aac atg gcc aaa caa gcc	830
Gln Trp Glu Lys Lys Leu Gln Thr Asp Ile Asn Met Ala Lys Gln Ala	
160 165 170	
tta ttc caa gcc ttg tca ctt gac caa cca tct tca ttg atc cct ccc	878
Leu Phe Gln Ala Leu Ser Leu Asp Gln Pro Ser Ser Leu Ile Pro Pro	
175 180 185	
gat cct gac tca cca aaa cct cat cat cat tct acc acc act tat gcc	926
Asp Pro Asp Ser Pro Lys Pro His His His Ser Thr Thr Thr Tyr Ala	
190 195 200 205	
tca agc aca gat aac atc tct aaa tta ctc cag aac tgg aca agc tca	974
Ser Ser Thr Asp Asn Ile Ser Lys Leu Leu Gln Asn Trp Thr Ser Ser	
210 215 220	
tca tcg tca aag cct aac act tca tca gtc tcc aac aac cgg agc tca	1022
Ser Ser Ser Lys Pro Asn Thr Ser Ser Val Ser Asn Asn Arg Ser Ser	
225 230 235	
agc ccc ggt gaa gga gga ctt ttt gat cat cac tct ttg ttc tca tcg	1070
Ser Pro Gly Glu Gly Gly Leu Phe Asp His His Ser Leu Phe Ser Ser	
240 245 250	
aat tca gaa tct gga tca gtt gat gag aag ctg aat ttg atg tcc gag	1118
Asn Ser Glu Ser Gly Ser Val Asp Glu Lys Leu Asn Leu Met Ser Glu	
255 260 265	
aca agc atg ttc aaa ggt gag agc aag cca gac ata gac atg gaa gct	1166
Thr Ser Met Phe Lys Gly Glu Ser Lys Pro Asp Ile Asp Met Glu Ala	
270 275 280 285	
aca cct act act act act act act act gat gat caa ggc tcg ttg tca	1214
Thr Pro Thr Thr Thr Thr Thr Thr Thr Asp Asp Gln Gly Ser Leu Ser	
290 295 300	
ttg atc gag aaa tgg ttg ttt gat gat caa ggc ttg gtt cag tgt gat	1262
Leu Ile Glu Lys Trp Leu Phe Asp Asp Gln Gly Leu Val Gln Cys Asp	
305 310 315	
gat agt caa gaa gat ctc atc gac gtg tct tta gag gag tta aaa taa	1310
Asp Ser Gln Glu Asp Leu Ile Asp Val Ser Leu Glu Glu Leu Lys *	
320 325 330	
tgataacaac agtcaagatt tgttctataa gaaaataaaa cgtatagaac aacgataaag	1370
ctagctagggt ttattaattt ttctttcttt tgtcttttct ctatgatctt tagttacatt	1430
ttattttact gtgtggcttg cttgtgggtca agtcgatgaa gatcaaaactg tgatatacta	1490
tttatatgta aagtactata aagttaagag tagttgaata aaaaaaaaaa aaaaaaaa	1547

<210> 6  
 <211> 332  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (13)...(115)  
 <223> Conserved domain

<400> 6  
 Met Gly Arg Pro Pro Cys Cys Glu Lys Ile Glu Val Lys Lys Gly Pro  
 1 5 10 15  
 Trp Thr Pro Glu Asp Ile Ile Leu Val Ser Tyr Ile Gln Gln His  
 20 25 30  
 Gly Pro Gly Asn Trp Arg Ser Val Pro Ala Asn Thr Gly Leu Leu Arg  
 35 40 45  
 Cys Ser Lys Ser Cys Arg Leu Arg Trp Thr Asn Tyr Leu Arg Pro Gly  
 50 55 60  
 Ile Lys Arg Gly Asn Phe Thr Gln Pro Glu Glu Lys Met Ile Ile His  
 65 70 75 80  
 Leu Gln Ala Leu Leu Gly Asn Arg Trp Ala Ala Ile Ala Ser Tyr Leu  
 85 90 95  
 Pro Gln Arg Thr Asp Asn Asp Ile Lys Asn Tyr Trp Asn Thr His Leu  
 100 105 110  
 Lys Lys Lys Leu Val Met Met Lys Phe Gln Asn Gly Ile Ile Asn Glu  
 115 120 125  
 Asn Lys Thr Asn Leu Ala Thr Asp Ile Ser Ser Cys Asn Asn Asn Asn  
 130 135 140  
 Asn Gly Cys Asn His Asn Lys Arg Thr Thr Asn Lys Gly Gln Trp Glu  
 145 150 155 160  
 Lys Lys Leu Gln Thr Asp Ile Asn Met Ala Lys Gln Ala Leu Phe Gln  
 165 170 175  
 Ala Leu Ser Leu Asp Gln Pro Ser Ser Leu Ile Pro Pro Asp Pro Asp  
 180 185 190  
 Ser Pro Lys Pro His His His Ser Thr Thr Tyr Ala Ser Ser Thr  
 195 200 205  
 Asp Asn Ile Ser Lys Leu Leu Gln Asn Trp Thr Ser Ser Ser Ser Ser  
 210 215 220  
 Lys Pro Asn Thr Ser Ser Val Ser Asn Asn Arg Ser Ser Ser Pro Gly  
 225 230 235 240  
 Glu Gly Gly Leu Phe Asp His His Ser Leu Phe Ser Ser Asn Ser Glu  
 245 250 255  
 Ser Gly Ser Val Asp Glu Lys Leu Asn Leu Met Ser Glu Thr Ser Met  
 260 265 270  
 Phe Lys Gly Glu Ser Lys Pro Asp Ile Asp Met Glu Ala Thr Pro Thr  
 275 280 285  
 Thr Thr Thr Thr Thr Asp Asp Gln Gly Ser Leu Ser Leu Ile Glu  
 290 295 300  
 Lys Trp Leu Phe Asp Asp Gln Gly Leu Val Gln Cys Asp Asp Ser Gln  
 305 310 315 320  
 Glu Asp Leu Ile Asp Val Ser Leu Glu Glu Leu Lys  
 325 330

<210> 7

&lt;211&gt; 1249

&lt;212&gt; DNA

<213> *Arabidopsis thaliana*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (82)...(918)

&lt;400&gt; 7

```

gaaattctta acaaacaatt ttcttcataa tattaattct caagatctta aagattatat 60
taatacgaag agaaaattca a atg ggt ctt gat gat tca tgc aac aca ggt 111
                               Met Gly Leu Asp Asp Ser Cys Asn Thr Gly
                               1           5           10

ctt gtt ctt ggt tta ggc ctc tca cca acg cct aat aat tac aat cat 159
Leu Val Leu Gly Leu Gly Leu Ser Pro Thr Pro Asn Asn Tyr Asn His
                               15           20           25

gcc atc aag aaa tct tcc tcc act gtg gac cat cgt ttc atc agg ctc 207
Ala Ile Lys Lys Ser Ser Ser Thr Val Asp His Arg Phe Ile Arg Leu
                               30           35           40

gat ccg tcg ttg act cta agc cta tcc ggt gag agc tac aag atc aag 255
Asp Pro Ser Leu Thr Leu Ser Leu Ser Gly Glu Ser Tyr Lys Ile Lys
                               45           50           55

act ggt gcc ggc gcc ggc gac caa att tgc cgg cag acc tcg tcc cac 303
Thr Gly Ala Gly Ala Gly Asp Gln Ile Cys Arg Gln Thr Ser Ser His
                               60           65           70

agc ggc atc tca tct ttc tcg agc gga agg gta aag aga gaa aga gaa 351
Ser Gly Ile Ser Ser Phe Ser Ser Gly Arg Val Lys Arg Glu Arg Glu
                               75           80           85           90

atc tcc ggc ggc gat gga gaa gaa gag gcg gag gag acg acg gag aga 399
Ile Ser Gly Gly Asp Gly Glu Glu Glu Ala Glu Glu Thr Thr Glu Arg
                               95           100           105

gtg gtg tgt tcg aga gtg agt gat gat cat gac gat gaa gaa ggt gtt 447
Val Val Cys Ser Arg Val Ser Asp Asp His Asp Asp Glu Glu Gly Val
                               110           115           120

agt gct cgt aaa aag ctt aga ctc act aaa caa caa tct gct ctt ctc 495
Ser Ala Arg Lys Lys Leu Arg Leu Thr Lys Gln Gln Ser Ala Leu Leu
                               125           130           135

gaa gat aac ttc aaa ctt cat agc acc ctt aat ccc aag caa aaa caa 543
Glu Asp Asn Phe Lys Leu His Ser Thr Leu Asn Pro Lys Gln Lys Gln
                               140           145           150

gct ctt gcg aga cag ctg aat cta agg cct aga caa gtt gaa gtg tgg 591
Ala Leu Ala Arg Gln Leu Asn Leu Arg Pro Arg Gln Val Glu Val Trp
                               155           160           165           170

ttc caa aac agg aga gct aga aca aaa cta aag caa aca gaa gtg gat 639
Phe Gln Asn Arg Arg Ala Arg Thr Lys Leu Lys Gln Thr Glu Val Asp
                               175           180           185

```



tgt gag ttt ttg aag aaa tgt tgc gag act tta acg gat gag aat aga 687  
 Cys Glu Phe Leu Lys Lys Cys Cys Glu Thr Leu Thr Asp Glu Asn Arg  
                   190                                  195                                  200

agg ctt caa aaa gag ctt caa gac ctt aag gct tta aaa ttg tct caa 735  
 Arg Leu Gln Lys Glu Leu Gln Asp Leu Lys Ala Leu Lys Leu Ser Gln  
                   205                                  210                                  215

ccg ttt tac atg cac atg ccg gcg gcg act ttg act atg tgc cct tct 783  
 Pro Phe Tyr Met His Met Pro Ala Ala Thr Leu Thr Met Cys Pro Ser  
                   220                                  225                                  230

tgt gag aga ctc ggc ggt ggt ggt gtc gga gga gat acg acg gcg gtt 831  
 Cys Glu Arg Leu Gly Gly Gly Gly Val Gly Gly Asp Thr Thr Ala Val  
                   235                                  240                                  245                                  250

gat gaa gaa acg gcg aaa gga gct ttc tcc atc gtc aca aag cct cgt 879  
 Asp Glu Glu Thr Ala Lys Gly Ala Phe Ser Ile Val Thr Lys Pro Arg  
                                   255                                  260                                  265

ttc tat aac cct ttc act aat cct tct gca gca tgt tag ttacttatta 928  
 Phe Tyr Asn Pro Phe Thr Asn Pro Ser Ala Ala Cys \*  
                   270                                  275

gttatttaaat tcttttttgtt gggttttttt ttgttttctta aatcaaatta ggaattagtt 988  
 agaagataaa tcccagggaa aaaatattac gttgaaattg ggggggaaatg gggatatagtc 1048  
 tttatagata agactcttca acgattccac tttatttttc ggtgggattg ttggttgatg 1108  
 aagaaaaaaaa aatagtttgt aattacaggt ttaaataatgt agagaaaaaaaa tgacgaatat 1168  
 gtattatctt gttttttttt ccttcgaata tgtattacgg taatataaat ttgcttgtaa 1228  
 aaataataaa tatattattt g 1249

<210> 8

<211> 278

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (121)...(182)

<223> Conserved domain

<400> 8

Met Gly Leu Asp Asp Ser Cys Asn Thr Gly Leu Val Leu Gly Leu Gly  
   1                  5                                  10                                  15  
 Leu Ser Pro Thr Pro Asn Asn Tyr Asn His Ala Ile Lys Lys Ser Ser  
                   20                                  25                                  30  
 Ser Thr Val Asp His Arg Phe Ile Arg Leu Asp Pro Ser Leu Thr Leu  
                   35                                  40                                  45  
 Ser Leu Ser Gly Glu Ser Tyr Lys Ile Lys Thr Gly Ala Gly Ala Gly  
   50                  55                                  60  
 Asp Gln Ile Cys Arg Gln Thr Ser Ser His Ser Gly Ile Ser Ser Phe  
   65                  70                                  75                                  80  
 Ser Ser Gly Arg Val Lys Arg Glu Arg Glu Ile Ser Gly Gly Asp Gly  
                   85                                  90                                  95  
 Glu Glu Glu Ala Glu Glu Thr Thr Glu Arg Val Val Cys Ser Arg Val  
                   100                                  105                                  110

Ser Asp Asp His Asp Asp Glu Glu Gly Val Ser Ala Arg Lys Lys Leu  
 115 120 125  
 Arg Leu Thr Lys Gln Gln Ser Ala Leu Leu Glu Asp Asn Phe Lys Leu  
 130 135 140  
 His Ser Thr Leu Asn Pro Lys Gln Lys Gln Ala Leu Ala Arg Gln Leu  
 145 150 155 160  
 Asn Leu Arg Pro Arg Gln Val Glu Val Trp Phe Gln Asn Arg Arg Ala  
 165 170 175  
 Arg Thr Lys Leu Lys Gln Thr Glu Val Asp Cys Glu Phe Leu Lys Lys  
 180 185 190  
 Cys Cys Glu Thr Leu Thr Asp Glu Asn Arg Arg Leu Gln Lys Glu Leu  
 195 200 205  
 Gln Asp Leu Lys Ala Leu Lys Leu Ser Gln Pro Phe Tyr Met His Met  
 210 215 220  
 Pro Ala Ala Thr Leu Thr Met Cys Pro Ser Cys Glu Arg Leu Gly Gly  
 225 230 235 240  
 Gly Gly Val Gly Gly Asp Thr Thr Ala Val Asp Glu Glu Thr Ala Lys  
 245 250 255  
 Gly Ala Phe Ser Ile Val Thr Lys Pro Arg Phe Tyr Asn Pro Phe Thr  
 260 265 270  
 Asn Pro Ser Ala Ala Cys  
 275

&lt;210&gt; 9

&lt;211&gt; 1087

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (104)...(952)

&lt;400&gt; 9

caatccctca atataaaata acaagtagaa ttgatctgcc tatatataag attttgagac 60  
 gaaataagat ctaaaccaca agaaagaaag taaacataaa agt atg gga agg tca 115  
 Met Gly Arg Ser  
 1

ccg tgc tgt gag aaa gct cac aca aac aaa gga gca tgg acg aaa gaa 163  
 Pro Cys Cys Glu Lys Ala His Thr Asn Lys Gly Ala Trp Thr Lys Glu  
 5 10 15 20

gag gac gag agg ctc gtc gcc tac att aaa gct cat gga gaa ggc tgc 211  
 Glu Asp Glu Arg Leu Val Ala Tyr Ile Lys Ala His Gly Glu Gly Cys  
 25 30 35

tgg aga tct ctc ccc aaa gcc gcc gga ctt ctt cgc tgt ggc aag agc 259  
 Trp Arg Ser Leu Pro Lys Ala Ala Gly Leu Leu Arg Cys Gly Lys Ser  
 40 45 50

tgc cgt ctc cgg tgg atc aac tat ctc cgg cct gac ctt aag cgt gga 307  
 Cys Arg Leu Arg Trp Ile Asn Tyr Leu Arg Pro Asp Leu Lys Arg Gly  
 55 60 65

aac ttc acc gag gaa gaa gac gaa ctc atc atc aag ctc cat agc ctt 355  
 Asn Phe Thr Glu Glu Glu Asp Glu Leu Ile Ile Lys Leu His Ser Leu

70	75	80	
ctt ggc aac aaa tgg tgc ctt att gcc ggg aga tta ccg gga aga aca			403
Leu Gly Asn Lys Trp Ser Leu Ile Ala Gly Arg Leu Pro Gly Arg Thr			
85	90	95	100
gat aac gag ata aag aac tat tgg aac acg cat ata cga aga aag ctt			451
Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr His Ile Arg Arg Lys Leu			
105	110	115	
ata aac aga ggg att gat cca acg agt cat aga cca atc caa gaa tca			499
Ile Asn Arg Gly Ile Asp Pro Thr Ser His Arg Pro Ile Gln Glu Ser			
120	125	130	
tca gct tct caa gat tct aaa cct aca caa cta gaa cca gtt acg agt			547
Ser Ala Ser Gln Asp Ser Lys Pro Thr Gln Leu Glu Pro Val Thr Ser			
135	140	145	
aat acc att aat atc tca ttc act tct gct cca aag gtc gaa acg ttc			595
Asn Thr Ile Asn Ile Ser Phe Thr Ser Ala Pro Lys Val Glu Thr Phe			
150	155	160	
cat gaa agt ata agc ttt ccg gga aaa tca gag aaa atc tca atg ctt			643
His Glu Ser Ile Ser Phe Pro Gly Lys Ser Glu Lys Ile Ser Met Leu			
165	170	175	180
acg ttc aaa gaa gaa aaa gat gag tgc cca gtt caa gaa aag ttc cca			691
Thr Phe Lys Glu Glu Lys Asp Glu Cys Pro Val Gln Glu Lys Phe Pro			
185	190	195	
gat ttg aat ctt gag ctc aga atc agt ctt cct gat gat gtt gat cgt			739
Asp Leu Asn Leu Glu Leu Arg Ile Ser Leu Pro Asp Asp Val Asp Arg			
200	205	210	
ctt caa ggg cat gga aag tca aca acg cca cgt tgt ttc aag tgc agc			787
Leu Gln Gly His Gly Lys Ser Thr Thr Pro Arg Cys Phe Lys Cys Ser			
215	220	225	
tta ggg atg ata aac ggc atg gag tgc aga tgc gga aga atg aga tgc			835
Leu Gly Met Ile Asn Gly Met Glu Cys Arg Cys Gly Arg Met Arg Cys			
230	235	240	
gat gta gtc gga ggt agc agc aag ggg agt gac atg agc aat gga ttt			883
Asp Val Val Gly Gly Ser Ser Lys Gly Ser Asp Met Ser Asn Gly Phe			
245	250	255	260
gat ttt tta ggg ttg gca aag aaa gag acc act tct ctt ttg ggc ttt			931
Asp Phe Leu Gly Leu Ala Lys Lys Glu Thr Thr Ser Leu Leu Gly Phe			
265	270	275	
cga agc ttg gag atg aaa taa tattgtcaaa ttttaggcgt aactgtacaa			982
Arg Ser Leu Glu Met Lys *			
280			
aacttttggc tagataattt gaaagtatat cttcaacttg tatgagaaat ttaactggtg			1042
aattataata tatagaattt gttttttaaa aaaaaaaaaa aaaaa			1087

<210> 10  
 <211> 282  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (13)...(116)  
 <223> Conserved domain

<400> 10  
 Met Gly Arg Ser Pro Cys Cys Glu Lys Ala His Thr Asn Lys Gly Ala  
 1 5 10 15  
 Trp Thr Lys Glu Glu Asp Glu Arg Leu Val Ala Tyr Ile Lys Ala His  
 20 25 30  
 Gly Glu Gly Cys Trp Arg Ser Leu Pro Lys Ala Ala Gly Leu Leu Arg  
 35 40 45  
 Cys Gly Lys Ser Cys Arg Leu Arg Trp Ile Asn Tyr Leu Arg Pro Asp  
 50 55 60  
 Leu Lys Arg Gly Asn Phe Thr Glu Glu Glu Asp Glu Leu Ile Ile Lys  
 65 70 75 80  
 Leu His Ser Leu Leu Gly Asn Lys Trp Ser Leu Ile Ala Gly Arg Leu  
 85 90 95  
 Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr His Ile  
 100 105 110  
 Arg Arg Lys Leu Ile Asn Arg Gly Ile Asp Pro Thr Ser His Arg Pro  
 115 120 125  
 Ile Gln Glu Ser Ser Ala Ser Gln Asp Ser Lys Pro Thr Gln Leu Glu  
 130 135 140  
 Pro Val Thr Ser Asn Thr Ile Asn Ile Ser Phe Thr Ser Ala Pro Lys  
 145 150 155 160  
 Val Glu Thr Phe His Glu Ser Ile Ser Phe Pro Gly Lys Ser Glu Lys  
 165 170 175  
 Ile Ser Met Leu Thr Phe Lys Glu Glu Lys Asp Glu Cys Pro Val Gln  
 180 185 190  
 Glu Lys Phe Pro Asp Leu Asn Leu Glu Leu Arg Ile Ser Leu Pro Asp  
 195 200 205  
 Asp Val Asp Arg Leu Gln Gly His Gly Lys Ser Thr Thr Pro Arg Cys  
 210 215 220  
 Phe Lys Cys Ser Leu Gly Met Ile Asn Gly Met Glu Cys Arg Cys Gly  
 225 230 235 240  
 Arg Met Arg Cys Asp Val Val Gly Gly Ser Ser Lys Gly Ser Asp Met  
 245 250 255  
 Ser Asn Gly Phe Asp Phe Leu Gly Leu Ala Lys Lys Glu Thr Thr Ser  
 260 265 270  
 Leu Leu Gly Phe Arg Ser Leu Glu Met Lys  
 275 280

<210> 11  
 <211> 1747  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (503)...(1534)

```

<400> 11
tgcaaaaaca ttttcttgtc tctcctctgc ccaaattttt tttctttcca ggaatatttc 60
ctagaaaaac ccaagcaaag cttaacccc ttctcctcc aaaagtagca tcttctcttc 120
tttctatttc tcttttcttc ttcttatctc tctctcgttt gtgaacgatt ccttaagaat 180
ataacaaaaa gcccttttct cctttcttca actttccggg aaaaatcttc acgcagcaag 240
gtttctctct cggtctctgc agtggttttc gggccttttg ttctttctat aaaaaaaaaa 300
ttcgcgctct ttaagaaaac tttttccacc tagagaagaa gaagagtatc actcttggtg 360
ttcaagtttc tctctttaat aaaaaatcca tctttattct ttgtcttctt tcttttttgc 420
tttccctaat ctctatgtta taaacacaca gagagaaaca aagtcacagt ctcgagtcaa 480
aaacagagaa taagaaagaa aa atg gaa gcg gag aag aaa atg gtt cta ccg 532
                               Met Glu Ala Glu Lys Lys Met Val Leu Pro
                               1             5             10

aga atc aaa ttc aca gag cac aaa acc aac acg aca aca atc gta tcg 580
Arg Ile Lys Phe Thr Glu His Lys Thr Asn Thr Thr Thr Ile Val Ser
                               15             20             25

gag tta acc aac act cac caa acc agg att ctt cgt atc tca gtc act 628
Glu Leu Thr Asn Thr His Gln Thr Arg Ile Leu Arg Ile Ser Val Thr
                               30             35             40

gac cca gac gct act gat tcc tcc agt gac gac gaa gaa gaa gaa cat 676
Asp Pro Asp Ala Thr Asp Ser Ser Ser Asp Asp Glu Glu Glu Glu His
                               45             50             55

caa cgc ttt gtc tct aaa cgc cgt cgt gtt aag aag ttt gtc aac gaa 724
Gln Arg Phe Val Ser Lys Arg Arg Arg Val Lys Lys Phe Val Asn Glu
                               60             65             70

gtc tat ctc gat tcc ggt gct gtt gtt act ggt agt tgt ggt caa atg 772
Val Tyr Leu Asp Ser Gly Ala Val Val Thr Gly Ser Cys Gly Gln Met
                               75             80             85             90

gag tcg aag aag aga caa aag aga gcg gtt aaa tcg gag tct act gtt 820
Glu Ser Lys Lys Arg Gln Lys Arg Ala Val Lys Ser Glu Ser Thr Val
                               95             100             105

tct ccg gtt gtt tca gcg acg acg act acg acg gga gag aag aag ttc 868
Ser Pro Val Val Ser Ala Thr Thr Thr Thr Gly Glu Lys Lys Phe
                               110             115             120

cga gga gtg aga cag cgt cca tgg gga aaa tgg gcg gcg gag ata aga 916
Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Trp Ala Ala Glu Ile Arg
                               125             130             135

gat ccg ttg aaa cgt gta cgg ctc tgg tta ggt act tac aac acg gcg 964
Asp Pro Leu Lys Arg Val Arg Leu Trp Leu Gly Thr Tyr Asn Thr Ala
                               140             145             150

gaa gaa gct gct atg gtt tac gat aac gcc gct att cag ctt cgt ggt 1012
Glu Glu Ala Ala Met Val Tyr Asp Asn Ala Ala Ile Gln Leu Arg Gly
                               155             160             165             170

ccc gac gct ctg act aat ttc tca gtc act ccg aca aca gcg acg gag 1060
Pro Asp Ala Leu Thr Asn Phe Ser Val Thr Pro Thr Thr Ala Thr Glu
                               175             180             185

```

aag aaa gcc cca cca ccg tct ccg gtg aag aag aag aag aag aaa aac 1108  
 Lys Lys Ala Pro Pro Pro Ser Pro Val Lys Lys Lys Lys Lys Lys Asn  
 190 195 200  
 aac aaa agc aaa aaa tcc gtt act gct tct tcc tcc atc agc aga agc 1156  
 Asn Lys Ser Lys Lys Ser Val Thr Ala Ser Ser Ser Ile Ser Arg Ser  
 205 210 215  
 agc agc aac gat tgt ctc tgc tct ccg gtg tct gtt ctc cga tct cct 1204  
 Ser Ser Asn Asp Cys Leu Cys Ser Pro Val Ser Val Leu Arg Ser Pro  
 220 225 230  
 ttc gcc gtc gac gaa ttc tcc ggc att tct tca tca cca gtc gcg gcc 1252  
 Phe Ala Val Asp Glu Phe Ser Gly Ile Ser Ser Ser Pro Val Ala Ala  
 235 240 245 250  
 gtt gta gtc aag gaa gag cca tcc atg aca acg gta tct gaa act ttc 1300  
 Val Val Val Lys Glu Glu Pro Ser Met Thr Thr Val Ser Glu Thr Phe  
 255 260 265  
 tct gat ttc tcg gcg ccc ttg ttc tca gat gat gac gtg ttc gat ttc 1348  
 Ser Asp Phe Ser Ala Pro Leu Phe Ser Asp Asp Asp Val Phe Asp Phe  
 270 275 280  
 cgg agc tca gtg gtt ccc gac tat ctc ggc ggc gat tta ttt ggg gaa 1396  
 Arg Ser Ser Val Val Pro Asp Tyr Leu Gly Gly Asp Leu Phe Gly Glu  
 285 290 295  
 gat cta ttc acg gcg gat atg tgt acg gat atg aac ttc gga ttc gat 1444  
 Asp Leu Phe Thr Ala Asp Met Cys Thr Asp Met Asn Phe Gly Phe Asp  
 300 305 310  
 ttc gga tcc gga tta tcc agc tgg cac atg gag gac cat ttt caa gat 1492  
 Phe Gly Ser Gly Leu Ser Ser Trp His Met Glu Asp His Phe Gln Asp  
 315 320 325 330  
 atc ggg gat cta ttc ggg tcg gat cct ctt tta gct gtt taa 1534  
 Ile Gly Asp Leu Phe Gly Ser Asp Pro Leu Leu Ala Val \*  
 335 340  
 taatatttta aataaataaa tagttatacc ggccgttact aaacggaacc ggagaaagtt 1594  
 ttgtataccg gtgacataaa atctcggtta tgttcgtaat ctttttttct ttgttatata 1654  
 taaaaatatg aatgaaactg aattaatgta agttaatggt gataattatt aacgttttaa 1714  
 gttttgaaaa aaaaaaaaaa aaaaaaaaaa aaa 1747

&lt;210&gt; 12

&lt;211&gt; 343

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (119)...(186)

&lt;223&gt; Conserved domain

&lt;400&gt; 12

Met Glu Ala Glu Lys Lys Met Val Leu Pro Arg Ile Lys Phe Thr Glu  
 1 5 10 15  
 His Lys Thr Asn Thr Thr Thr Ile Val Ser Glu Leu Thr Asn Thr His  
 20 25 30  
 Gln Thr Arg Ile Leu Arg Ile Ser Val Thr Asp Pro Asp Ala Thr Asp  
 35 40 45  
 Ser Ser Ser Asp Asp Glu Glu Glu His Gln Arg Phe Val Ser Lys  
 50 55 60  
 Arg Arg Arg Val Lys Lys Phe Val Asn Glu Val Tyr Leu Asp Ser Gly  
 65 70 75 80  
 Ala Val Val Thr Gly Ser Cys Gly Gln Met Glu Ser Lys Lys Arg Gln  
 85 90 95  
 Lys Arg Ala Val Lys Ser Glu Ser Thr Val Ser Pro Val Val Ser Ala  
 100 105 110  
 Thr Thr Thr Thr Thr Gly Glu Lys Lys Phe Arg Gly Val Arg Gln Arg  
 115 120 125  
 Pro Trp Gly Lys Trp Ala Ala Glu Ile Arg Asp Pro Leu Lys Arg Val  
 130 135 140  
 Arg Leu Trp Leu Gly Thr Tyr Asn Thr Ala Glu Glu Ala Ala Met Val  
 145 150 155 160  
 Tyr Asp Asn Ala Ala Ile Gln Leu Arg Gly Pro Asp Ala Leu Thr Asn  
 165 170 175  
 Phe Ser Val Thr Pro Thr Thr Ala Thr Glu Lys Lys Ala Pro Pro Pro  
 180 185 190  
 Ser Pro Val Lys Lys Lys Lys Lys Lys Asn Asn Lys Ser Lys Lys Ser  
 195 200 205  
 Val Thr Ala Ser Ser Ser Ile Ser Arg Ser Ser Ser Asn Asp Cys Leu  
 210 215 220  
 Cys Ser Pro Val Ser Val Leu Arg Ser Pro Phe Ala Val Asp Glu Phe  
 225 230 235 240  
 Ser Gly Ile Ser Ser Ser Pro Val Ala Ala Val Val Val Lys Glu Glu  
 245 250 255  
 Pro Ser Met Thr Thr Val Ser Glu Thr Phe Ser Asp Phe Ser Ala Pro  
 260 265 270  
 Leu Phe Ser Asp Asp Asp Val Phe Asp Phe Arg Ser Ser Val Val Pro  
 275 280 285  
 Asp Tyr Leu Gly Gly Asp Leu Phe Gly Glu Asp Leu Phe Thr Ala Asp  
 290 295 300  
 Met Cys Thr Asp Met Asn Phe Gly Phe Asp Phe Gly Ser Gly Leu Ser  
 305 310 315 320  
 Ser Trp His Met Glu Asp His Phe Gln Asp Ile Gly Asp Leu Phe Gly  
 325 330 335  
 Ser Asp Pro Leu Leu Ala Val  
 340

&lt;210&gt; 13

&lt;211&gt; 609

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)... (609)

&lt;400&gt; 13

atg gct gag aac aac aac aac aac ggc gac aac atg aac aac gac aac 48

Met	Ala	Glu	Asn	Asn	Asn	Asn	Asn	Gly	Asp	Asn	Met	Asn	Asn	Asp	Asn		
1				5				10						15			
cac	cag	caa	cca	ccg	tcg	tac	tcg	cag	ctg	ccg	ccg	atg	gca	tca	tcc	96	
His	Gln	Gln	Pro	Pro	Ser	Tyr	Ser	Gln	Leu	Pro	Pro	Met	Ala	Ser	Ser		
			20					25					30				
aac	cct	cag	tta	cgt	aat	tac	tgg	att	gag	cag	atg	gaa	acc	gtc	tcg	144	
Asn	Pro	Gln	Leu	Arg	Asn	Tyr	Trp	Ile	Glu	Gln	Met	Glu	Thr	Val	Ser		
		35				40					45						
gat	ttc	aaa	aac	cgt	cag	ctt	cca	ttg	gct	cga	att	aag	aag	atc	atg	192	
Asp	Phe	Lys	Asn	Arg	Gln	Leu	Pro	Leu	Ala	Arg	Ile	Lys	Lys	Ile	Met		
	50					55				60							
aag	gct	gat	cca	gat	gtg	cac	atg	gtc	tcc	gca	gag	gct	ccg	atc	atc	240	
Lys	Ala	Asp	Pro	Asp	Val	His	Met	Val	Ser	Ala	Glu	Ala	Pro	Ile	Ile		
	65				70					75				80			
ttc	gca	aag	gct	tgc	gaa	atg	ttc	atc	gtt	gat	ctc	acg	atg	cgg	tcg	288	
Phe	Ala	Lys	Ala	Cys	Glu	Met	Phe	Ile	Val	Asp	Leu	Thr	Met	Arg	Ser		
				85				90						95			
tgg	ctc	aaa	gcc	gag	gag	aac	aaa	cgc	cac	acg	ctt	cag	aaa	tcg	gat	336	
Trp	Leu	Lys	Ala	Glu	Glu	Asn	Lys	Arg	His	Thr	Leu	Gln	Lys	Ser	Asp		
			100					105					110				
atc	tcc	aac	gca	gtg	gct	agc	tct	ttc	acc	tac	gat	ttc	ctt	ctt	gat	384	
Ile	Ser	Asn	Ala	Val	Ala	Ser	Ser	Phe	Thr	Tyr	Asp	Phe	Leu	Leu	Asp		
		115					120					125					
gtt	gtc	cct	aag	gac	gag	tct	atc	gcc	acc	gct	gat	cct	ggc	ttt	gtg	432	
Val	Val	Pro	Lys	Asp	Glu	Ser	Ile	Ala	Thr	Ala	Asp	Pro	Gly	Phe	Val		
	130					135					140						
gct	atg	cca	cat	cct	gac	ggc	gga	gga	gta	ccg	caa	tat	tat	tat	cca	480	
Ala	Met	Pro	His	Pro	Asp	Gly	Gly	Gly	Val	Pro	Gln	Tyr	Tyr	Tyr	Pro		
	145				150					155					160		
ccg	gga	gtg	gtg	atg	gga	act	cct	atg	gtt	ggc	agt	gga	atg	tac	gcg	528	
Pro	Gly	Val	Val	Met	Gly	Thr	Pro	Met	Val	Gly	Ser	Gly	Met	Tyr	Ala		
				165				170						175			
cca	tcg	cag	gcg	tgg	cca	gca	gcg	gct	ggc	gac	ggg	gag	gat	gat	gct	576	
Pro	Ser	Gln	Ala	Trp	Pro	Ala	Ala	Ala	Gly	Asp	Gly	Glu	Asp	Asp	Ala		
		180					185						190				
gag	gat	aat	gga	gga	aac	ggc	ggc	gga	aat	tga						609	
Glu	Asp	Asn	Gly	Gly	Asn	Gly	Gly	Gly	Asn	*							
		195				200											

&lt;210&gt; 14

&lt;211&gt; 202

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana



<220>  
 <221> DOMAIN  
 <222> (70)...(133)  
 <223> Conserved domain

<400> 14  
 Met Ala Glu Asn Asn Asn Asn Asn Gly Asp Asn Met Asn Asn Asp Asn  
 1 5 10 15  
 His Gln Gln Pro Pro Ser Tyr Ser Gln Leu Pro Pro Met Ala Ser Ser  
 20 25 30  
 Asn Pro Gln Leu Arg Asn Tyr Trp Ile Glu Gln Met Glu Thr Val Ser  
 35 40 45  
 Asp Phe Lys Asn Arg Gln Leu Pro Leu Ala Arg Ile Lys Lys Ile Met  
 50 55 60  
 Lys Ala Asp Pro Asp Val His Met Val Ser Ala Glu Ala Pro Ile Ile  
 65 70 75 80  
 Phe Ala Lys Ala Cys Glu Met Phe Ile Val Asp Leu Thr Met Arg Ser  
 85 90 95  
 Trp Leu Lys Ala Glu Glu Asn Lys Arg His Thr Leu Gln Lys Ser Asp  
 100 105 110  
 Ile Ser Asn Ala Val Ala Ser Ser Phe Thr Tyr Asp Phe Leu Leu Asp  
 115 120 125  
 Val Val Pro Lys Asp Glu Ser Ile Ala Thr Ala Asp Pro Gly Phe Val  
 130 135 140  
 Ala Met Pro His Pro Asp Gly Gly Gly Val Pro Gln Tyr Tyr Tyr Pro  
 145 150 155 160  
 Pro Gly Val Val Met Gly Thr Pro Met Val Gly Ser Gly Met Tyr Ala  
 165 170 175  
 Pro Ser Gln Ala Trp Pro Ala Ala Ala Gly Asp Gly Glu Asp Asp Ala  
 180 185 190  
 Glu Asp Asn Gly Gly Asn Gly Gly Gly Asn  
 195 200

<210> 15  
 <211> 724  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (20)...(694)

<400> 15  
 catcttatcc aaagaaaaa atg aat cca ttt tac tct aca ttc cca gac tcg 52  
 Met Asn Pro Phe Tyr Ser Thr Phe Pro Asp Ser  
 1 5 10  
 ttt ctc tca atc tcc gat cat aga tct ccg gtt tca gac agt agt gag 100  
 Phe Leu Ser Ile Ser Asp His Arg Ser Pro Val Ser Asp Ser Ser Glu  
 15 20 25  
 tgt tca cca aag tta gct tca agt tgt cca aag aaa cga gct ggg agg 148  
 Cys Ser Pro Lys Leu Ala Ser Ser Cys Pro Lys Lys Arg Ala Gly Arg  
 30 35 40  
 aag aag ttt cgt gag aca cgt cat ccg att tac aga gga gtt cgt cag 196

```

Lys Lys Phe Arg Glu Thr Arg His Pro Ile Tyr Arg Gly Val Arg Gln
 45                               50                               55

agg aat tct ggt aaa tgg gtt tgt gaa gtt aga gag cct aat aag aaa 244
Arg Asn Ser Gly Lys Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys
 60                               65                               70                               75

tct agg att tgg tta ggt act ttt ccg acg gtt gaa atg gct gct cgt 292
Ser Arg Ile Trp Leu Gly Thr Phe Pro Thr Val Glu Met Ala Ala Arg
 80                               85                               90

gct cat gat gtt gct gct tta gct ctt cgt ggt cgc tct gct tgt ctc 340
Ala His Asp Val Ala Ala Leu Ala Leu Arg Gly Arg Ser Ala Cys Leu
 95                               100                               105

aat ttc gct gat tct gct tgg cgg ctt cgt att cct gag act act tgt 388
Asn Phe Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys
110                               115                               120

cct aag gag att cag aaa gct gcg tct gaa gct gca atg gcg ttt cag 436
Pro Lys Glu Ile Gln Lys Ala Ala Ser Glu Ala Ala Met Ala Phe Gln
125                               130                               135

aat gag act acg acg gag gga tct aaa act gcg gcg gag gca gag gag 484
Asn Glu Thr Thr Thr Glu Gly Ser Lys Thr Ala Ala Glu Ala Glu Glu
140                               145                               150                               155

gcg gca ggg gag ggg gtg agg gag ggg gag agg agg gcg gag gag cag 532
Ala Ala Gly Glu Gly Val Arg Glu Gly Glu Arg Arg Ala Glu Glu Gln
160                               165                               170

aat ggt ggt gtg ttt tat atg gat gat gag gcg ctt ttg ggg atg ccc 580
Asn Gly Gly Val Phe Tyr Met Asp Asp Glu Ala Leu Leu Gly Met Pro
175                               180                               185

aac ttt ttt gag aat atg gcg gag ggg atg ctt ttg ccg ccg ccg gaa 628
Asn Phe Phe Glu Asn Met Ala Glu Gly Met Leu Leu Pro Pro Pro Glu
190                               195                               200

gtt ggc tgg aat cat aac gac ttt gac gga gtg ggt gac gtg tca ctc 676
Val Gly Trp Asn His Asn Asp Phe Asp Gly Val Gly Asp Val Ser Leu
205                               210                               215

tgg agt ttt gac gag taa ttttttggt ctttttctgg ataataagtt 724
Trp Ser Phe Asp Glu *
220

```

&lt;210&gt; 16

&lt;211&gt; 224

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (51)...(118)

&lt;223&gt; Conserved domain

&lt;400&gt; 16

```

Met Asn Pro Phe Tyr Ser Thr Phe Pro Asp Ser Phe Leu Ser Ile Ser
 1           5           10           15
Asp His Arg Ser Pro Val Ser Asp Ser Ser Glu Cys Ser Pro Lys Leu
          20           25           30
Ala Ser Ser Cys Pro Lys Lys Arg Ala Gly Arg Lys Lys Phe Arg Glu
        35           40           45
Thr Arg His Pro Ile Tyr Arg Gly Val Arg Gln Arg Asn Ser Gly Lys
 50           55           60
Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu
65           70           75           80
Gly Thr Phe Pro Thr Val Glu Met Ala Ala Arg Ala His Asp Val Ala
          85           90           95
Ala Leu Ala Leu Arg Gly Arg Ser Ala Cys Leu Asn Phe Ala Asp Ser
        100          105          110
Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys Pro Lys Glu Ile Gln
       115          120          125
Lys Ala Ala Ser Glu Ala Ala Met Ala Phe Gln Asn Glu Thr Thr Thr
      130          135          140
Glu Gly Ser Lys Thr Ala Ala Glu Ala Glu Glu Ala Ala Gly Glu Gly
     145          150          155          160
Val Arg Glu Gly Glu Arg Arg Ala Glu Glu Gln Asn Gly Gly Val Phe
          165          170          175
Tyr Met Asp Asp Glu Ala Leu Leu Gly Met Pro Asn Phe Phe Glu Asn
        180          185          190
Met Ala Glu Gly Met Leu Leu Pro Pro Pro Glu Val Gly Trp Asn His
       195          200          205
Asn Asp Phe Asp Gly Val Gly Asp Val Ser Leu Trp Ser Phe Asp Glu
      210          215          220

```

&lt;210&gt; 17

&lt;211&gt; 653

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (38)...(607)

&lt;400&gt; 17

```

caaatcagaa aatatagagt ttgaaggaaa ctaaaag atg gta cat tcg agg aag 55
                               Met Val His Ser Arg Lys
                               1           5

ttc cga ggt gtc cgc cag cga caa tgg ggt tct tgg gtc tct gag att 103
Phe Arg Gly Val Arg Gln Arg Gln Trp Gly Ser Trp Val Ser Glu Ile
          10           15           20

cgc cat cct cta ttg aag aga aga gtg tgg ctt gga act ttc gaa acg 151
Arg His Pro Leu Leu Lys Arg Arg Val Trp Leu Gly Thr Phe Glu Thr
        25           30           35

gca gaa gcg gct gca aga gca tac gac caa gcg gct ctt cta atg aac 199
Ala Glu Ala Ala Ala Arg Ala Tyr Asp Gln Ala Ala Leu Leu Met Asn

```

40 45 50  
 ggc caa aac gct aag acc aat ttc cct gtc gta aaa tca gag gaa ggc 247  
 Gly Gln Asn Ala Lys Thr Asn Phe Pro Val Val Lys Ser Glu Glu Gly  
 55 60 65 70  
 tcc gat cac gtt aaa gat gtt aac tct ccg ttg atg tca cca aag tca 295  
 Ser Asp His Val Lys Asp Val Asn Ser Pro Leu Met Ser Pro Lys Ser  
 75 80 85  
 tta tct gag ctt ttg aac gct aag cta agg aag agc tgc aaa gac cta 343  
 Leu Ser Glu Leu Leu Asn Ala Lys Leu Arg Lys Ser Cys Lys Asp Leu  
 90 95 100  
 acg cct tct ttg acg tgt ctc cgt ctt gat act gac agt tcc cac att 391  
 Thr Pro Ser Leu Thr Cys Leu Arg Leu Asp Thr Asp Ser Ser His Ile  
 105 110 115  
 gga gtt tgg cag aaa cgg gcc ggg tcg aaa aca agt ccg act tgg gtc 439  
 Gly Val Trp Gln Lys Arg Ala Gly Ser Lys Thr Ser Pro Thr Trp Val  
 120 125 130  
 atg cgc ctc gaa ctt ggg aac gta gtc aac gaa agt gcg gtt gac tta 487  
 Met Arg Leu Glu Leu Gly Asn Val Val Asn Glu Ser Ala Val Asp Leu  
 135 140 145 150  
 ggg ttg act acg atg aac aaa caa aac gtt gag aaa gaa gaa gaa gaa 535  
 Gly Leu Thr Thr Met Asn Lys Gln Asn Val Glu Lys Glu Glu Glu Glu  
 155 160 165  
 gaa gaa gct att att agt gat gag gat cag tta gct atg gag atg atc 583  
 Glu Glu Ala Ile Ile Ser Asp Glu Asp Gln Leu Ala Met Glu Met Ile  
 170 175 180  
 gag gag ttg ctg aat tgg agt tga cttttgactt taacttggtg caagtccaca 637  
 Glu Glu Leu Leu Asn Trp Ser \*  
 185  
 aggggtaagg gttttc 653  
 <210> 18  
 <211> 768  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (58)...(657)  
 <400> 18  
 attactcatc atcaagttcc tactttctct ctgacaaaca tcacagagta agtaaga atg 60  
 Met  
 1  
 gta cag acg aag aag ttc aga ggt gtc agg caa cgc cat tgg ggt tct 108  
 Val Gln Thr Lys Lys Phe Arg Gly Val Arg Gln Arg His Trp Gly Ser  
 5 10 15

tgg gtc gct gag att cgt cat cct ctc ttg aaa cgg agg att tgg cta 156  
 Trp Val Ala Glu Ile Arg His Pro Leu Leu Lys Arg Arg Ile Trp Leu  
 20 25 30

ggg acg ttc gag acc gca gag gag gca gca aga gca tac gac gag gcc 204  
 Gly Thr Phe Glu Thr Ala Glu Glu Ala Ala Arg Ala Tyr Asp Glu Ala  
 35 40 45

gcc gtt tta atg agc ggc cgc aac gcc aaa acc aac ttt ccc ctc aac 252  
 Ala Val Leu Met Ser Gly Arg Asn Ala Lys Thr Asn Phe Pro Leu Asn  
 50 55 60 65

aac aac aac acc gga gaa act tcc gag ggc aaa acc gat att tca gct 300  
 Asn Asn Asn Thr Gly Glu Thr Ser Glu Gly Lys Thr Asp Ile Ser Ala  
 70 75 80

tcg tcc aca atg tca tcc tca aca tca tct tca tcg ctc tct tcc atc 348  
 Ser Ser Thr Met Ser Ser Ser Thr Ser Ser Ser Ser Leu Ser Ser Ile  
 85 90 95

ctc agc gcc aaa ctg agg aaa tgc tgc aag tct cct tcc cca tcc ctc 396  
 Leu Ser Ala Lys Leu Arg Lys Cys Cys Lys Ser Pro Ser Pro Ser Leu  
 100 105 110

acc tgc ctc cgt ctt gac aca gcc agc tcc cat atc ggc gtc tgg cag 444  
 Thr Cys Leu Arg Leu Asp Thr Ala Ser Ser His Ile Gly Val Trp Gln  
 115 120 125

aaa cgg gcc ggt tca aag tct gac tcc agc tgg gtc atg acg gtg gag 492  
 Lys Arg Ala Gly Ser Lys Ser Asp Ser Ser Trp Val Met Thr Val Glu  
 130 135 140 145

cta ggt ccc gca agc tcc tcc caa gag act act agt aaa gct tca caa 540  
 Leu Gly Pro Ala Ser Ser Ser Gln Glu Thr Thr Ser Lys Ala Ser Gln  
 150 155 160

gac gct att ctt gct ccg acc act gaa gtt gaa att ggt ggc agc aga 588  
 Asp Ala Ile Leu Ala Pro Thr Thr Glu Val Glu Ile Gly Gly Ser Arg  
 165 170 175

gaa gaa gta ttg gat gag gaa gaa aag gtt gct ttg caa atg ata gag 636  
 Glu Glu Val Leu Asp Glu Glu Glu Lys Val Ala Leu Gln Met Ile Glu  
 180 185 190

gag ctt ctc aat aca aac taa atcttatttg cttatatata tgtacctatt 687  
 Glu Leu Leu Asn Thr Asn \*  
 195

ttcattgctg atttacagcc aaaataatca attataccgt gtattttata gatgttttat 747  
 attaaaagggt tgtagatat a 768

<210> 19  
 <211> 189  
 <212> PRT  
 <213> Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (4)...(71)

&lt;223&gt; Conserved domain

&lt;400&gt; 19

```

Met Val His Ser Arg Lys Phe Arg Gly Val Arg Gln Arg Gln Trp Gly
 1           5           10           15
Ser Trp Val Ser Glu Ile Arg His Pro Leu Leu Lys Arg Arg Val Trp
          20           25           30
Leu Gly Thr Phe Glu Thr Ala Glu Ala Ala Arg Ala Tyr Asp Gln
      35           40           45
Ala Ala Leu Leu Met Asn Gly Gln Asn Ala Lys Thr Asn Phe Pro Val
      50           55           60
Val Lys Ser Glu Glu Gly Ser Asp His Val Lys Asp Val Asn Ser Pro
65           70           75           80
Leu Met Ser Pro Lys Ser Leu Ser Glu Leu Leu Asn Ala Lys Leu Arg
          85           90           95
Lys Ser Cys Lys Asp Leu Thr Pro Ser Leu Thr Cys Leu Arg Leu Asp
          100          105          110
Thr Asp Ser Ser His Ile Gly Val Trp Gln Lys Arg Ala Gly Ser Lys
          115          120          125
Thr Ser Pro Thr Trp Val Met Arg Leu Glu Leu Gly Asn Val Val Asn
          130          135          140
Glu Ser Ala Val Asp Leu Gly Leu Thr Thr Met Asn Lys Gln Asn Val
145          150          155          160
Glu Lys Glu Glu Glu Glu Glu Ala Ile Ile Ser Asp Glu Asp Gln
          165          170          175
Leu Ala Met Glu Met Ile Glu Glu Leu Leu Asn Trp Ser
          180          185

```

&lt;210&gt; 20

&lt;211&gt; 199

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (4)...(71)

&lt;223&gt; Conserved domain

&lt;400&gt; 20

```

Met Val Gln Thr Lys Lys Phe Arg Gly Val Arg Gln Arg His Trp Gly
 1           5           10           15
Ser Trp Val Ala Glu Ile Arg His Pro Leu Leu Lys Arg Arg Ile Trp
          20           25           30
Leu Gly Thr Phe Glu Thr Ala Glu Ala Ala Arg Ala Tyr Asp Glu
      35           40           45
Ala Ala Val Leu Met Ser Gly Arg Asn Ala Lys Thr Asn Phe Pro Leu
      50           55           60
Asn Asn Asn Asn Thr Gly Glu Thr Ser Glu Gly Lys Thr Asp Ile Ser
65           70           75           80
Ala Ser Ser Thr Met Ser Ser Ser Thr Ser Ser Ser Ser Leu Ser Ser
          85           90           95
Ile Leu Ser Ala Lys Leu Arg Lys Cys Cys Lys Ser Pro Ser Pro Ser
          100          105          110

```

Leu Thr Cys Leu Arg Leu Asp Thr Ala Ser Ser His Ile Gly Val Trp  
           115                          120                          125  
 Gln Lys Arg Ala Gly Ser Lys Ser Asp Ser Ser Trp Val Met Thr Val  
       130                          135                          140  
 Glu Leu Gly Pro Ala Ser Ser Ser Gln Glu Thr Thr Ser Lys Ala Ser  
       145                          150                          155                          160  
 Gln Asp Ala Ile Leu Ala Pro Thr Thr Glu Val Glu Ile Gly Gly Ser  
                           165                          170                          175  
 Arg Glu Glu Val Leu Asp Glu Glu Glu Lys Val Ala Leu Gln Met Ile  
                           180                          185                          190  
 Glu Glu Leu Leu Asn Thr Asn  
                           195

<210> 21

<211> 1011

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (213)...(833)

<400> 21

aaagttattg atagtttctg ttacttatta atttttaagg ttatgtgtat tattaccaat 60  
 tggaggacta tatagtcgca agtctcaacc ctataaaaga aaacattcgt cgatcatctt 120  
 cccgcctcga gtatctctct ctctctctct ctctctctgtt ttctttattg attgcataga 180  
 caaaaatata cacatacaca acagaaagaa ag atg gag acg acg atg aag aag 233  
   Met Glu Thr Thr Met Lys Lys  
   1                          5

aaa ggg aga gtg aaa gcg aca ata acg tca cag aaa gaa gaa gaa gga 281  
 Lys Gly Arg Val Lys Ala Thr Ile Thr Ser Gln Lys Glu Glu Glu Gly  
       10                          15                          20

aca gtg aga aaa gga cct tgg act atg gaa gaa gat ttc atc ctc ttt 329  
 Thr Val Arg Lys Gly Pro Trp Thr Met Glu Glu Asp Phe Ile Leu Phe  
       25                          30                          35

aat tac atc ctt aat cat ggt gaa ggt ctt tgg aac tct gtc gcc aaa 377  
 Asn Tyr Ile Leu Asn His Gly Glu Gly Leu Trp Asn Ser Val Ala Lys  
       40                          45                          50                          55

gcc tct ggt cta aaa cgt act gga aaa agt tgt cgg ctc cgg tgg ctg 425  
 Ala Ser Gly Leu Lys Arg Thr Gly Lys Ser Cys Arg Leu Arg Trp Leu  
                           60                          65                          70

aac tat ctc cga cca gat gtg cgg cga ggg aac ata acc gaa gaa gaa 473  
 Asn Tyr Leu Arg Pro Asp Val Arg Arg Gly Asn Ile Thr Glu Glu Glu  
                           75                          80                          85

cag ctt ttg atc att cag ctt cat gct aag ctt gga aac agg tgg tcg 521  
 Gln Leu Leu Ile Ile Gln Leu His Ala Lys Leu Gly Asn Arg Trp Ser  
       90                          95                          100

aag att gcg aag cat ctt ccg gga aga acg gac aac gag ata aag aac 569  
 Lys Ile Ala Lys His Leu Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn

105                      110                      115  
 ttc tgg agg aca aag att cag aga cac atg aaa gtg tca tcg gaa aat 617  
 Phe Trp Arg Thr Lys Ile Gln Arg His Met Lys Val Ser Ser Glu Asn  
 120                      125                      130                      135  
 atg atg aat cat caa cat cat tgt tcg gga aac tca cag agc tcg ggg 665  
 Met Met Asn His Gln His His Cys Ser Gly Asn Ser Gln Ser Ser Gly  
 140                      145                      150  
 atg acg acg caa ggc agc tcc ggc aaa gcc ata gac acg gct gag agc 713  
 Met Thr Thr Gln Gly Ser Ser Gly Lys Ala Ile Asp Thr Ala Glu Ser  
 155                      160                      165  
 ttc tct cag gcg aag acg acg acg ttt aat gtg gtg gaa caa cag tca 761  
 Phe Ser Gln Ala Lys Thr Thr Thr Phe Asn Val Val Glu Gln Gln Ser  
 170                      175                      180  
 aac gag aat tac tgg aac gtt gaa gat ctg tgg ccc gtc cac ttg ctt 809  
 Asn Glu Asn Tyr Trp Asn Val Glu Asp Leu Trp Pro Val His Leu Leu  
 185                      190                      195  
 aat ggt gac cac cat gtg att taa gatatatata tagacctcct atacatttat 863  
 Asn Gly Asp His His Val Ile \*  
 200                      205  
 atgccccagc tgggtttttt tgtatggtac gttatttggt ttttctattg ctgaaatgct 923  
 gttgcattta atttacatac gaaaagtgc ttaaatcatt aaatcttcaa tacatatgga 983  
 ggtggtgttt gagtaaaaaa aaaaaaaa 1011

&lt;210&gt; 22

&lt;211&gt; 206

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (26)...(130)

&lt;223&gt; Conserved domain

&lt;400&gt; 22

Met Glu Thr Thr Met Lys Lys Lys Gly Arg Val Lys Ala Thr Ile Thr  
 1                      5                      10                      15  
 Ser Gln Lys Glu Glu Glu Gly Thr Val Arg Lys Gly Pro Trp Thr Met  
 20                      25                      30  
 Glu Glu Asp Phe Ile Leu Phe Asn Tyr Ile Leu Asn His Gly Glu Gly  
 35                      40                      45  
 Leu Trp Asn Ser Val Ala Lys Ala Ser Gly Leu Lys Arg Thr Gly Lys  
 50                      55                      60  
 Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp Val Arg Arg  
 65                      70                      75                      80  
 Gly Asn Ile Thr Glu Glu Gln Leu Leu Ile Ile Gln Leu His Ala  
 85                      90                      95  
 Lys Leu Gly Asn Arg Trp Ser Lys Ile Ala Lys His Leu Pro Gly Arg  
 100                      105                      110  
 Thr Asp Asn Glu Ile Lys Asn Phe Trp Arg Thr Lys Ile Gln Arg His  
 115                      120                      125



Met Lys Val Ser Ser Glu Asn Met Met Asn His Gln His His Cys Ser  
 130 135 140  
 Gly Asn Ser Gln Ser Ser Gly Met Thr Thr Gln Gly Ser Ser Gly Lys  
 145 150 155 160  
 Ala Ile Asp Thr Ala Glu Ser Phe Ser Gln Ala Lys Thr Thr Thr Phe  
 165 170 175  
 Asn Val Val Glu Gln Gln Ser Asn Glu Asn Tyr Trp Asn Val Glu Asp  
 180 185 190  
 Leu Trp Pro Val His Leu Leu Asn Gly Asp His His Val Ile  
 195 200 205

<210> 23  
 <211> 1072  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (87)...(977)

<400> 23  
 aaaagataac gtagccatta acttttatgc attaactcct tcatttcttt ttgtgctcgt 60  
 ttggttgaga gagagagagg ctatat atg atg ttc caa caa gat tac cct cat 113  
 Met Met Phe Gln Gln Asp Tyr Pro His  
 1 5

ggc ttt tca ctc gtg gaa aca tcc tta agt tac gaa atg ttg gat tac 161  
 Gly Phe Ser Leu Val Glu Thr Ser Leu Ser Tyr Glu Met Leu Asp Tyr  
 10 15 20 25

ttt caa aac atc gtc gtt tcg aac tct gaa gac gtg gcg tca cag caa 209  
 Phe Gln Asn Ile Val Ser Asn Ser Glu Asp Val Ala Ser Gln Gln  
 30 35 40

aat tcc att tcg tcc tct tct tat tca tca gcg aca ctc tcc tgc tcc 257  
 Asn Ser Ile Ser Ser Ser Ser Tyr Ser Ser Ala Thr Leu Ser Cys Ser  
 45 50 55

ata aca gag caa aaa tct cac tta act gaa aag tta tct cct cta cga 305  
 Ile Thr Glu Gln Lys Ser His Leu Thr Glu Lys Leu Ser Pro Leu Arg  
 60 65 70

gaa aga tat ggt tgc ggt gac ttt ctg tcg cgg aag agg aga agg aga 353  
 Glu Arg Tyr Gly Cys Gly Asp Phe Leu Ser Arg Lys Arg Arg Arg Arg  
 75 80 85

agt gaa aaa acg att gta gat aaa gag aat caa agg atg aat cac att 401  
 Ser Glu Lys Thr Ile Val Asp Lys Glu Asn Gln Arg Met Asn His Ile  
 90 95 100 105

gcc gtc gag cgt aac cgg aga aaa cag atg aat cat ttt ctg tct atc 449  
 Ala Val Glu Arg Asn Arg Arg Lys Gln Met Asn His Phe Leu Ser Ile  
 110 115 120

ctc aag tct atg atg cct ctc tct tat tct caa cct aat gac caa gca 497  
 Leu Lys Ser Met Met Pro Leu Ser Tyr Ser Gln Pro Asn Asp Gln Ala

125	130	135	
tca atc ata gaa ggg acc att agc tat ctg aag aag cta gaa caa cgt			545
Ser Ile Ile Glu Gly Thr Ile Ser Tyr Leu Lys Lys Leu Glu Gln Arg			
140	145	150	
ctc caa tct ctc gaa gcc caa tta aaa gct act aaa ctc aat caa tca			593
Leu Gln Ser Leu Glu Ala Gln Leu Lys Ala Thr Lys Leu Asn Gln Ser			
155	160	165	
cca aat ata ttt tcc gac ttc ttc atg ttc cct caa tac tcc acc gcc			641
Pro Asn Ile Phe Ser Asp Phe Phe Met Phe Pro Gln Tyr Ser Thr Ala			
170	175	180	185
act gcc act gcc acc gcc act gcc tcc tca tcc tcc tcg agc cac cac			689
Thr Ala Thr Ala Thr Ala Thr Ala Ser Ser Ser Ser Ser Ser His His			
190	195	200	
cat cac aag cga cta gag gtg gtt gct gac gtg gag gtt aca atg gta			737
His His Lys Arg Leu Glu Val Val Ala Asp Val Glu Val Thr Met Val			
205	210	215	
gaa aga cat gcc aac att aaa gtg tta acg aag aca cag cca aga ttg			785
Glu Arg His Ala Asn Ile Lys Val Leu Thr Lys Thr Gln Pro Arg Leu			
220	225	230	
ctc ttc aag att atc aat gag ttt aac tct tta ggt tta agt act ctt			833
Leu Phe Lys Ile Ile Asn Glu Phe Asn Ser Leu Gly Leu Ser Thr Leu			
235	240	245	
cat ctc aac ctc aca act tcc aaa gac atg tct ctc ttc act ttt agc			881
His Leu Asn Leu Thr Thr Ser Lys Asp Met Ser Leu Phe Thr Phe Ser			
250	255	260	265
gtc aag gta gag gca gat tgt caa ttg acg cct tct ggt aat gag gtc			929
Val Lys Val Glu Ala Asp Cys Gln Leu Thr Pro Ser Gly Asn Glu Val			
270	275	280	
gca aat acg gtg cat gaa gtc gtt aga aga gtt cac aag gaa cgt tga			977
Ala Asn Thr Val His Glu Val Val Arg Arg Val His Lys Glu Arg *			
285	290	295	
atattgttta catactagct aactttgaaa ttctatttta ttgtataaac aatctcttat			1037
gtgtgtaatt tacatatata caattaatta acatt			1072
<210> 24			
<211> 296			
<212> PRT			
<213> Arabidopsis thaliana			
<220>			
<221> DOMAIN			
<222> (92)...(161)			
<223> Conserved domain			
<400> 24			
Met Met Phe Gln Gln Asp Tyr Pro His Gly Phe Ser Leu Val Glu Thr			

```

1           5           10           15
Ser Leu Ser Tyr Glu Met Leu Asp Tyr Phe Gln Asn Ile Val Val Ser
20           25           30
Asn Ser Glu Asp Val Ala Ser Gln Gln Asn Ser Ile Ser Ser Ser
35           40           45
Tyr Ser Ser Ala Thr Leu Ser Cys Ser Ile Thr Glu Gln Lys Ser His
50           55           60
Leu Thr Glu Lys Leu Ser Pro Leu Arg Glu Arg Tyr Gly Cys Gly Asp
65           70           75           80
Phe Leu Ser Arg Lys Arg Arg Arg Arg Ser Glu Lys Thr Ile Val Asp
85           90           95
Lys Glu Asn Gln Arg Met Asn His Ile Ala Val Glu Arg Asn Arg Arg
100          105          110
Lys Gln Met Asn His Phe Leu Ser Ile Leu Lys Ser Met Met Pro Leu
115          120          125
Ser Tyr Ser Gln Pro Asn Asp Gln Ala Ser Ile Ile Glu Gly Thr Ile
130          135          140
Ser Tyr Leu Lys Lys Leu Glu Gln Arg Leu Gln Ser Leu Glu Ala Gln
145          150          155          160
Leu Lys Ala Thr Lys Leu Asn Gln Ser Pro Asn Ile Phe Ser Asp Phe
165          170          175
Phe Met Phe Pro Gln Tyr Ser Thr Ala Thr Ala Thr Ala Thr
180          185          190
Ala Ser Ser Ser Ser Ser His His His His Lys Arg Leu Glu Val
195          200          205
Val Ala Asp Val Glu Val Thr Met Val Glu Arg His Ala Asn Ile Lys
210          215          220
Val Leu Thr Lys Thr Gln Pro Arg Leu Leu Phe Lys Ile Ile Asn Glu
225          230          235          240
Phe Asn Ser Leu Gly Leu Ser Thr Leu His Leu Asn Leu Thr Thr Ser
245          250          255
Lys Asp Met Ser Leu Phe Thr Phe Ser Val Lys Val Glu Ala Asp Cys
260          265          270
Gln Leu Thr Pro Ser Gly Asn Glu Val Ala Asn Thr Val His Glu Val
275          280          285
Val Arg Arg Val His Lys Glu Arg
290          295

```

&lt;210&gt; 25

&lt;211&gt; 2209

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (186)...(1793)

&lt;400&gt; 25

```

ctttgaatgt gactcaggat agagagaggg cgagacagat aaagagagct tgaggaagga 60
tttaagcaga tctgcaagaa acactctctc agtggttact ttttttccgt agaatcaaac 120
cggcggcgac agtgacggcg gcgaaaagcg gcggtgttgg ccgcaaaaac gatgcaaaac 180
cagct atg gtc att acc gct aac gat tta tca aaa tgg gaa aat ttt cct 230
Met Val Ile Thr Ala Asn Asp Leu Ser Lys Trp Glu Asn Phe Pro
1           5           10           15

```

```

aaa gga ctt aag gtt ctt ctt ctc ctc aac ggc tgt gac agc gac gga 278

```

Lys Gly Leu Lys Val Leu Leu Leu Leu Asn Gly Cys Asp Ser Asp Gly	
20 25 30	
gat ggc tcc tca gcc gcc gag act cga tca gag ctc gaa tca atg gac	326
Asp Gly Ser Ser Ala Ala Glu Thr Arg Ser Glu Leu Glu Ser Met Asp	
35 40 45	
tat atc gtt act aca ttc acc gat gaa act gaa gca ctc tct gcg gtt	374
Tyr Ile Val Thr Thr Phe Thr Asp Glu Thr Glu Ala Leu Ser Ala Val	
50 55 60	
gtc aag aac ccg gag agc ttc cac att gcc atc gtc gag gtg aat atg	422
Val Lys Asn Pro Glu Ser Phe His Ile Ala Ile Val Glu Val Asn Met	
65 70 75	
agc gct gag agt gag agt ttc aag ttt ctt gag gct gcc aaa gac gtt	470
Ser Ala Glu Ser Glu Ser Phe Lys Phe Leu Glu Ala Ala Lys Asp Val	
80 85 90 95	
ctt cct act ata atg att tca acc gat cat tgc atc act act aca atg	518
Leu Pro Thr Ile Met Ile Ser Thr Asp His Cys Ile Thr Thr Met	
100 105 110	
aaa tgc ata gcg ctt ggt gca gtt gag ttc cta caa aaa ccg ctc tca	566
Lys Cys Ile Ala Leu Gly Ala Val Glu Phe Leu Gln Lys Pro Leu Ser	
115 120 125	
ccg gag aaa tta aag aac att tgg cag cat gtt gtt cat aag gca ttc	614
Pro Glu Lys Leu Lys Asn Ile Trp Gln His Val Val His Lys Ala Phe	
130 135 140	
aat gat ggt gga agt aat gtt tcg ata tca ctt aag cca gtg aaa gaa	662
Asn Asp Gly Gly Ser Asn Val Ser Ile Ser Leu Lys Pro Val Lys Glu	
145 150 155	
tcc gtt gtc tcg atg ctt cat ctt gag acc gac atg aca atc gag gag	710
Ser Val Val Ser Met Leu His Leu Glu Thr Asp Met Thr Ile Glu Glu	
160 165 170 175	
aaa gat cca gcg cca tca aca ccg caa ttg aaa caa gat tca cgg tta	758
Lys Asp Pro Ala Pro Ser Thr Pro Gln Leu Lys Gln Asp Ser Arg Leu	
180 185 190	
cta gac ggt gat tgc caa gag aac ata aat ttc tcg atg gaa aat gta	806
Leu Asp Gly Asp Cys Gln Glu Asn Ile Asn Phe Ser Met Glu Asn Val	
195 200 205	
aat tcc tcg acc gag aaa gat aac atg gaa gat cat caa gac atc ggt	854
Asn Ser Ser Thr Glu Lys Asp Asn Met Glu Asp His Gln Asp Ile Gly	
210 215 220	
gaa tct aaa tca gtc gac act aca aac cgc aaa tta gat gac gac aaa	902
Glu Ser Lys Ser Val Asp Thr Thr Asn Arg Lys Leu Asp Asp Asp Lys	
225 230 235	
gtg gtt gtc aaa gaa gag aga gga gac agt gaa aaa gaa gaa ggt	950
Val Val Val Lys Glu Glu Arg Gly Asp Ser Glu Lys Glu Glu Glu Gly	

240	245	250	255	
gaa acc gga gat ctc ata agc gag aag aca gat tca gtt gat att cat				998
Glu Thr Gly Asp Leu Ile Ser Glu Lys Thr Asp Ser Val Asp Ile His				
	260	265	270	
aag aaa gaa gat gag act aaa ccg att aat aaa tca tcc ggg atc aag				1046
Lys Lys Glu Asp Glu Thr Lys Pro Ile Asn Lys Ser Ser Gly Ile Lys				
	275	280	285	
aac gtg tct ggt aac aaa act agt cga aag aag gtg gat tgg aca cca				1094
Asn Val Ser Gly Asn Lys Thr Ser Arg Lys Lys Val Asp Trp Thr Pro				
	290	295	300	
gag ctg cac aag aag ttt gtg caa gca gtt gag caa ctc ggc gtt gat				1142
Glu Leu His Lys Lys Phe Val Gln Ala Val Glu Gln Leu Gly Val Asp				
	305	310	315	
caa gcg ata ccc tcg cgg att ctt gag ttg atg aaa gta ggc acc tta				1190
Gln Ala Ile Pro Ser Arg Ile Leu Glu Leu Met Lys Val Gly Thr Leu				
	320	325	330	335
aca aga cac aat gta gct agt cac ctt cag aaa ttt cgg cag cat agg				1238
Thr Arg His Asn Val Ala Ser His Leu Gln Lys Phe Arg Gln His Arg				
	340	345	350	
aag aat att ctt cca aag gat gat cat aac cat aga tgg ata caa tct				1286
Lys Asn Ile Leu Pro Lys Asp Asp His Asn His Arg Trp Ile Gln Ser				
	355	360	365	
aga gag aac cat aga cca aat caa cgt aat tat aac gtt ttt caa cag				1334
Arg Glu Asn His Arg Pro Asn Gln Arg Asn Tyr Asn Val Phe Gln Gln				
	370	375	380	
caa cac cgt ccc gtg atg gct tat ccc gtt tgg ggt ctt ccc ggt gtt				1382
Gln His Arg Pro Val Met Ala Tyr Pro Val Trp Gly Leu Pro Gly Val				
	385	390	395	
tat ccg cca gga gcg att cca cct ttg tgg cca ccg ccg ctg cag tcc				1430
Tyr Pro Pro Gly Ala Ile Pro Pro Leu Trp Pro Pro Pro Leu Gln Ser				
	400	405	410	415
att ggt caa cca cct ccg tgg cat tgg aaa cca cct tat cca acg gtg				1478
Ile Gly Gln Pro Pro Pro Trp His Trp Lys Pro Pro Tyr Pro Thr Val				
	420	425	430	
agc ggt aat gca tgg ggt tgt ccg gtt gga ccg cct gtg acc gga tca				1526
Ser Gly Asn Ala Trp Gly Cys Pro Val Gly Pro Pro Val Thr Gly Ser				
	435	440	445	
tat att act cct tcg aat act acc gcc ggc gga ttt caa tat ccc aac				1574
Tyr Ile Thr Pro Ser Asn Thr Thr Ala Gly Gly Phe Gln Tyr Pro Asn				
	450	455	460	
gga gct gaa acc ggc ttc aaa ata atg ccg gcg agt cag ccg gac gag				1622
Gly Ala Glu Thr Gly Phe Lys Ile Met Pro Ala Ser Gln Pro Asp Glu				
	465	470	475	

gaa atg tta gat caa gtg gtt aaa gaa gcg ata agc aaa ccg tgg ctg 1670  
 Glu Met Leu Asp Gln Val Val Lys Glu Ala Ile Ser Lys Pro Trp Leu  
 480 485 490 495

ccg cta ccg ctc ggg cta aaa ccg ccg tcc gcg gag agc gtt ttg gct 1718  
 Pro Leu Pro Leu Gly Leu Lys Pro Pro Ser Ala Glu Ser Val Leu Ala  
 500 505 510

gag cta acg cgt caa ggc atc tca gcc gtc cct tct tct tct tgt cta 1766  
 Glu Leu Thr Arg Gln Gly Ile Ser Ala Val Pro Ser Ser Ser Cys Leu  
 515 520 525

atc aac ggc tct cat cgt ctc cgc tga cgtgtccatg atgaaacaga 1813  
 Ile Asn Gly Ser His Arg Leu Arg \*  
 530 535

accgcagat atgatgacgt catgcgtggc gtacttgtcc gtgtctgacg atccgaagcg 1873  
 ggaactgatt cggatcctga gagttgtttt ttaattgttg tttttaattt tttgttgtca 1933  
 taaaaacatt tgtctctatt ggtattatta aagaagatcg gacgggtggag atagaggcaa 1993  
 cgatgttagt ctgaggacca caagtggact ctgcggctgt aagtagatct cggacacggg 2053  
 attttgacaa gtcggaatct gacgtgtgtg cgtactcgt tttgtgggtt gtcgagaatt 2113  
 taatagttga cactgtttta ggtggatatg agatatgatg atgataataa taaaaaggta 2173  
 caactcaaag attcctgat gaaaaaaaa aaaaaa 2209

<210> 26

<211> 535

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (301)...(349)

<223> Conserved domain

<400> 26

Met Val Ile Thr Ala Asn Asp Leu Ser Lys Trp Glu Asn Phe Pro Lys  
 1 5 10 15  
 Gly Leu Lys Val Leu Leu Leu Asn Gly Cys Asp Ser Asp Gly Asp  
 20 25 30  
 Gly Ser Ser Ala Ala Glu Thr Arg Ser Glu Leu Glu Ser Met Asp Tyr  
 35 40 45  
 Ile Val Thr Thr Phe Thr Asp Glu Thr Glu Ala Leu Ser Ala Val Val  
 50 55 60  
 Lys Asn Pro Glu Ser Phe His Ile Ala Ile Val Glu Val Asn Met Ser  
 65 70 75 80  
 Ala Glu Ser Glu Ser Phe Lys Phe Leu Glu Ala Ala Lys Asp Val Leu  
 85 90 95  
 Pro Thr Ile Met Ile Ser Thr Asp His Cys Ile Thr Thr Thr Met Lys  
 100 105 110  
 Cys Ile Ala Leu Gly Ala Val Glu Phe Leu Gln Lys Pro Leu Ser Pro  
 115 120 125  
 Glu Lys Leu Lys Asn Ile Trp Gln His Val Val His Lys Ala Phe Asn  
 130 135 140  
 Asp Gly Gly Ser Asn Val Ser Ile Ser Leu Lys Pro Val Lys Glu Ser  
 145 150 155 160  
 Val Val Ser Met Leu His Leu Glu Thr Asp Met Thr Ile Glu Glu Lys

				165						170				175		
Asp	Pro	Ala	Pro	Ser	Thr	Pro	Gln	Leu	Lys	Gln	Asp	Ser	Arg	Leu	Leu	
			180					185					190			
Asp	Gly	Asp	Cys	Gln	Glu	Asn	Ile	Asn	Phe	Ser	Met	Glu	Asn	Val	Asn	
		195				200					205					
Ser	Ser	Thr	Glu	Lys	Asp	Asn	Met	Glu	Asp	His	Gln	Asp	Ile	Gly	Glu	
		210				215					220					
Ser	Lys	Ser	Val	Asp	Thr	Thr	Asn	Arg	Lys	Leu	Asp	Asp	Asp	Lys	Val	
225					230					235					240	
Val	Val	Lys	Glu	Glu	Arg	Gly	Asp	Ser	Glu	Lys	Glu	Glu	Glu	Gly	Glu	
			245						250					255		
Thr	Gly	Asp	Leu	Ile	Ser	Glu	Lys	Thr	Asp	Ser	Val	Asp	Ile	His	Lys	
			260					265					270			
Lys	Glu	Asp	Glu	Thr	Lys	Pro	Ile	Asn	Lys	Ser	Ser	Gly	Ile	Lys	Asn	
		275					280					285				
Val	Ser	Gly	Asn	Lys	Thr	Ser	Arg	Lys	Lys	Val	Asp	Trp	Thr	Pro	Glu	
		290				295					300					
Leu	His	Lys	Lys	Phe	Val	Gln	Ala	Val	Glu	Gln	Leu	Gly	Val	Asp	Gln	
305					310					315					320	
Ala	Ile	Pro	Ser	Arg	Ile	Leu	Glu	Leu	Met	Lys	Val	Gly	Thr	Leu	Thr	
			325						330					335		
Arg	His	Asn	Val	Ala	Ser	His	Leu	Gln	Lys	Phe	Arg	Gln	His	Arg	Lys	
		340						345					350			
Asn	Ile	Leu	Pro	Lys	Asp	Asp	His	Asn	His	Arg	Trp	Ile	Gln	Ser	Arg	
		355				360						365				
Glu	Asn	His	Arg	Pro	Asn	Gln	Arg	Asn	Tyr	Asn	Val	Phe	Gln	Gln	Gln	
		370				375					380					
His	Arg	Pro	Val	Met	Ala	Tyr	Pro	Val	Trp	Gly	Leu	Pro	Gly	Val	Tyr	
385					390					395					400	
Pro	Pro	Gly	Ala	Ile	Pro	Pro	Leu	Trp	Pro	Pro	Pro	Leu	Gln	Ser	Ile	
			405						410					415		
Gly	Gln	Pro	Pro	Pro	Trp	His	Trp	Lys	Pro	Pro	Tyr	Pro	Thr	Val	Ser	
			420					425					430			
Gly	Asn	Ala	Trp	Gly	Cys	Pro	Val	Gly	Pro	Pro	Val	Thr	Gly	Ser	Tyr	
		435				440					445					
Ile	Thr	Pro	Ser	Asn	Thr	Thr	Ala	Gly	Gly	Phe	Gln	Tyr	Pro	Asn	Gly	
		450				455					460					
Ala	Glu	Thr	Gly	Phe	Lys	Ile	Met	Pro	Ala	Ser	Gln	Pro	Asp	Glu	Glu	
465					470					475					480	
Met	Leu	Asp	Gln	Val	Val	Lys	Glu	Ala	Ile	Ser	Lys	Pro	Trp	Leu	Pro	
			485						490					495	</	

```
<210> 27
<211> 890
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>  
<221> CDS  
<222> (108) ... (806)
```

&lt;400&gt; 27

cattcaaaaa catcatatat atacacaaac acactttgat acaacaaaaa aaaacagaac 60  
 acaacaaaa acacattgta acattagttt aagcattaag cttcttt atg tcg aat 116  
 Met Ser Asn  
 1

aat aat aat tct ccg acc acc gtg aat caa gaa acg acg acg tct cgt 164  
 Asn Asn Asn Ser Pro Thr Thr Val Asn Gln Glu Thr Thr Thr Ser Arg  
 5 10 15

gaa gtc tca atc aca ttg cct act gat caa tct cct caa acc tca cca 212  
 Glu Val Ser Ile Thr Leu Pro Thr Asp Gln Ser Pro Gln Thr Ser Pro  
 20 25 30 35

gga tca tct tct tct cct tca ccg aga cct tcc ggt gga tca ccg gcg 260  
 Gly Ser Ser Ser Ser Pro Ser Pro Arg Pro Ser Gly Gly Ser Pro Ala  
 40 45 50

aga aga acg gcg act gga tta tcc ggc aag cac tct att ttc agg ggg 308  
 Arg Arg Thr Ala Thr Gly Leu Ser Gly Lys His Ser Ile Phe Arg Gly  
 55 60 65

att cga cta cgt aac gga aaa tgg gta tcg gag att aga gag cca cgt 356  
 Ile Arg Leu Arg Asn Gly Lys Trp Val Ser Glu Ile Arg Glu Pro Arg  
 70 75 80

aaa acg aca aga att tgg ctc ggg act tat ccg gta ccg gag atg gct 404  
 Lys Thr Thr Arg Ile Trp Leu Gly Thr Tyr Pro Val Pro Glu Met Ala  
 85 90 95

gcc gcc gct tac gac gtg gct gcg tta gct tta aaa gga ccc gac gcc 452  
 Ala Ala Ala Tyr Asp Val Ala Ala Leu Ala Lys Gly Pro Asp Ala  
 100 105 110 115

gtt ttg aat ttt cct ggt tta gct ttg act tac gtg gct ccg gtt tca 500  
 Val Leu Asn Phe Pro Gly Leu Ala Leu Thr Tyr Val Ala Pro Val Ser  
 120 125 130

aac tct gct gcg gat ata aga gcg gct gct agt aga gca gcg gag atg 548  
 Asn Ser Ala Ala Asp Ile Arg Ala Ala Ala Ser Arg Ala Ala Glu Met  
 135 140 145

aag caa ccg gat cag ggt ggg gat gag aag gta ttg gaa ccg gtt caa 596  
 Lys Gln Pro Asp Gln Gly Gly Asp Glu Lys Val Leu Glu Pro Val Gln  
 150 155 160

ccc ggc aaa gag gaa gaa tta gaa gaa gtg tcg tgt aac tcg tgt tcg 644  
 Pro Gly Lys Glu Glu Glu Leu Glu Glu Val Ser Cys Asn Ser Cys Ser  
 165 170 175

ttg gag ttt atg gat gag gaa gcg atg ttg aat atg ccg act ttg ttg 692  
 Leu Glu Phe Met Asp Glu Glu Ala Met Leu Asn Met Pro Thr Leu Leu  
 180 185 190 195

acg gag atg gct gaa ggg atg ttg atg agt cca ccg aga atg atg ata 740  
 Thr Glu Met Ala Glu Gly Met Leu Met Ser Pro Pro Arg Met Met Ile



200 205 210

cat ccg acg atg gaa gat gat tcg ccg gag aat cat gaa gga gat aat 788  
 His Pro Thr Met Glu Asp Asp Ser Pro Glu Asn His Glu Gly Asp Asn  
 215 220 225

ctt tgg agt tat aaa tga atccattgaa gctgctctct tttttattgt 836  
 Leu Trp Ser Tyr Lys \*  
 230

tttccggtcg aatgagattt tccccctttt tttttttctt tttgggtcgc tggtt 890

<210> 28

<211> 232

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (62)...(128)

<223> Conserved domain

<400> 28

Met Ser Asn Asn Asn Asn Ser Pro Thr Thr Val Asn Gln Glu Thr Thr  
 1 5 10 15  
 Thr Ser Arg Glu Val Ser Ile Thr Leu Pro Thr Asp Gln Ser Pro Gln  
 20 25 30  
 Thr Ser Pro Gly Ser Ser Ser Ser Pro Ser Pro Arg Pro Ser Gly Gly  
 35 40 45  
 Ser Pro Ala Arg Arg Thr Ala Thr Gly Leu Ser Gly Lys His Ser Ile  
 50 55 60  
 Phe Arg Gly Ile Arg Leu Arg Asn Gly Lys Trp Val Ser Glu Ile Arg  
 65 70 75 80  
 Glu Pro Arg Lys Thr Thr Arg Ile Trp Leu Gly Thr Tyr Pro Val Pro  
 85 90 95  
 Glu Met Ala Ala Ala Ala Tyr Asp Val Ala Ala Leu Ala Leu Lys Gly  
 100 105 110  
 Pro Asp Ala Val Leu Asn Phe Pro Gly Leu Ala Leu Thr Tyr Val Ala  
 115 120 125  
 Pro Val Ser Asn Ser Ala Ala Asp Ile Arg Ala Ala Ala Ser Arg Ala  
 130 135 140  
 Ala Glu Met Lys Gln Pro Asp Gln Gly Gly Asp Glu Lys Val Leu Glu  
 145 150 155 160  
 Pro Val Gln Pro Gly Lys Glu Glu Glu Leu Glu Glu Val Ser Cys Asn  
 165 170 175  
 Ser Cys Ser Leu Glu Phe Met Asp Glu Glu Ala Met Leu Asn Met Pro  
 180 185 190  
 Thr Leu Leu Thr Glu Met Ala Glu Gly Met Leu Met Ser Pro Pro Arg  
 195 200 205  
 Met Met Ile His Pro Thr Met Glu Asp Asp Ser Pro Glu Asn His Glu  
 210 215 220  
 Gly Asp Asn Leu Trp Ser Tyr Lys  
 225 230

<210> 29

<211> 720

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(720)

&lt;400&gt; 29

atg ggg aat aga aga gca cca tgc tgt gac aaa agc caa gtg aag aga	48
Met Gly Asn Arg Arg Ala Pro Cys Cys Asp Lys Ser Gln Val Lys Arg	
1 5 10 15	
ggg cca tgg agt gat gaa gaa agt gaa aga ctc aga tct ttt atc ctc	96
Gly Pro Trp Ser Asp Glu Glu Ser Glu Arg Leu Arg Ser Phe Ile Leu	
20 25 30	
aaa aat ggt cat caa aat tgg cga tct ctt ccc aaa ctc gct gga ttg	144
Lys Asn Gly His Gln Asn Trp Arg Ser Leu Pro Lys Leu Ala Gly Leu	
35 40 45	
atg aga tgc gga aag agt tgt cgt cta aga tgg ata aac tat ctg aga	192
Met Arg Cys Gly Lys Ser Cys Arg Leu Arg Trp Ile Asn Tyr Leu Arg	
50 55 60	
cca ggt ctc aaa cga ggc aac ttc acc aag gag gag gaa gat acc att	240
Pro Gly Leu Lys Arg Gly Asn Phe Thr Lys Glu Glu Glu Asp Thr Ile	
65 70 75 80	
atc cac ctt cac caa gct tat gga aac aag tgg tct aag atc gca tcc	288
Ile His Leu His Gln Ala Tyr Gly Asn Lys Trp Ser Lys Ile Ala Ser	
85 90 95	
aac ttc ccg gga agg aca gac aac gag atc aag aat gta tgg aac act	336
Asn Phe Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Val Trp Asn Thr	
100 105 110	
cat ctc aag aag cgg ttg gtg aag agg agc att tca tca tca tca tcc	384
His Leu Lys Lys Arg Leu Val Lys Arg Ser Ile Ser Ser Ser Ser	
115 120 125	
gat gtt acc aat cat tca gtg tct tct acc tct tct tcc tct tcc tca	432
Asp Val Thr Asn His Ser Val Ser Ser Thr Ser Ser Ser Ser Ser	
130 135 140	
atc tcg tca gtc ttg cag gac gtt att att aag agt gag agg cct aac	480
Ile Ser Ser Val Leu Gln Asp Val Ile Ile Lys Ser Glu Arg Pro Asn	
145 150 155 160	
cag gaa gag gag ttt ggg gaa atc ttg gtg gag caa atg gca tgt gga	528
Gln Glu Glu Glu Phe Gly Glu Ile Leu Val Glu Gln Met Ala Cys Gly	
165 170 175	
ttt gag gtg gat gca cca caa tca cta gaa tgt ctc ttt gac gat agc	576
Phe Glu Val Asp Ala Pro Gln Ser Leu Glu Cys Leu Phe Asp Asp Ser	
180 185 190	
cag gtt cct cct cct ata tct aaa cca gac tca cta caa acc cat ggg	624

Gln Val Pro Pro Ile Ser Lys Pro Asp Ser Leu Gln Thr His Gly  
 195 200 205

aag tca tca gat cac gaa ttt tgg agc cga ctg att gaa cca ggg ttc 672  
 Lys Ser Ser Asp His Glu Phe Trp Ser Arg Leu Ile Glu Pro Gly Phe  
 210 215 220

gat gat tac aat gag tgg ctc att ttc ttg gat aac caa act tgc tag 720  
 Asp Asp Tyr Asn Glu Trp Leu Ile Phe Leu Asp Asn Gln Thr Cys \*  
 225 230 235

<210> 30  
 <211> 239  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (15)...(118)  
 <223> Conserved domain

<400> 30  
 Met Gly Asn Arg Arg Ala Pro Cys Cys Asp Lys Ser Gln Val Lys Arg  
 1 5 10 15  
 Gly Pro Trp Ser Asp Glu Glu Ser Glu Arg Leu Arg Ser Phe Ile Leu  
 20 25 30  
 Lys Asn Gly His Gln Asn Trp Arg Ser Leu Pro Lys Leu Ala Gly Leu  
 35 40 45  
 Met Arg Cys Gly Lys Ser Cys Arg Leu Arg Trp Ile Asn Tyr Leu Arg  
 50 55 60  
 Pro Gly Leu Lys Arg Gly Asn Phe Thr Lys Glu Glu Asp Thr Ile  
 65 70 75 80  
 Ile His Leu His Gln Ala Tyr Gly Asn Lys Trp Ser Lys Ile Ala Ser  
 85 90 95  
 Asn Phe Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Val Trp Asn Thr  
 100 105 110  
 His Leu Lys Lys Arg Leu Val Lys Arg Ser Ile Ser Ser Ser Ser  
 115 120 125  
 Asp Val Thr Asn His Ser Val Ser Ser Thr Ser Ser Ser Ser Ser  
 130 135 140  
 Ile Ser Ser Val Leu Gln Asp Val Ile Ile Lys Ser Glu Arg Pro Asn  
 145 150 155 160  
 Gln Glu Glu Glu Phe Gly Glu Ile Leu Val Glu Gln Met Ala Cys Gly  
 165 170 175  
 Phe Glu Val Asp Ala Pro Gln Ser Leu Glu Cys Leu Phe Asp Asp Ser  
 180 185 190  
 Gln Val Pro Pro Pro Ile Ser Lys Pro Asp Ser Leu Gln Thr His Gly  
 195 200 205  
 Lys Ser Ser Asp His Glu Phe Trp Ser Arg Leu Ile Glu Pro Gly Phe  
 210 215 220  
 Asp Asp Tyr Asn Glu Trp Leu Ile Phe Leu Asp Asn Gln Thr Cys  
 225 230 235

<210> 31

&lt;211&gt; 1144

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (25)...(1104)

&lt;400&gt; 31

```

cgctcgacctc ccaacactaa ctcc atg ttt ata acg gaa aaa caa gtg tgg      51
Met Phe Ile Thr Glu Lys Gln Val Trp
      1              5

atg gat gag atc gtc gca aga aga gct tct tct tct tgg gac ttc cct      99
Met Asp Glu Ile Val Ala Arg Arg Ala Ser Ser Ser Trp Asp Phe Pro
      10              15              20              25

ttc aac gac att aat att cat cag cat cat cat cgt cac tgc aac aca      147
Phe Asn Asp Ile Asn Ile His Gln His His His Arg His Cys Asn Thr
      30              35              40

agt cat gag ttt gaa atc ttg aag agt cct ctt gga gat gta gcg gtt      195
Ser His Glu Phe Glu Ile Leu Lys Ser Pro Leu Gly Asp Val Ala Val
      45              50              55

cac gaa gaa gag agt aat aat aat aac cct aat ttc agt aac agc gag      243
His Glu Glu Glu Ser Asn Asn Asn Asn Pro Asn Phe Ser Asn Ser Glu
      60              65              70

agt ggt aag aag gag aca aca gat agt ggt cag tct tgg tcc tcg tcg      291
Ser Gly Lys Lys Glu Thr Thr Asp Ser Gly Gln Ser Trp Ser Ser Ser
      75              80              85

tct tca aaa cca tcg gtc ttg ggg aga gga cat tgg aga cca gct gaa      339
Ser Ser Lys Pro Ser Val Leu Gly Arg Gly His Trp Arg Pro Ala Glu
      90              95              100              105

gat gtt aaa ctc aaa gag ctt gtc tcc att tac ggc cca caa aac tgg      387
Asp Val Lys Leu Lys Glu Leu Val Ser Ile Tyr Gly Pro Gln Asn Trp
      110              115              120

aac ctc ata gct gaa aag ctt caa gga aga tct ggg aag agc tgt aga      435
Asn Leu Ile Ala Glu Lys Leu Gln Gly Arg Ser Gly Lys Ser Cys Arg
      125              130              135

cta cga tgg ttt aac caa ttg gac ccg agg ata aac cga aga gct ttc      483
Leu Arg Trp Phe Asn Gln Leu Asp Pro Arg Ile Asn Arg Arg Ala Phe
      140              145              150

aca gaa gaa gaa gag gag agg ctg atg caa gca cat agg ctt tat ggt      531
Thr Glu Glu Glu Glu Arg Leu Met Gln Ala His Arg Leu Tyr Gly
      155              160              165

aac aaa tgg gca atg att gcg agg ctt ttc cct ggt aga act gat aat      579
Asn Lys Trp Ala Met Ile Ala Arg Leu Phe Pro Gly Arg Thr Asp Asn
      170              175              180              185

```

```

tca gtg aag aac cat tgg cat gtt gtc atg gct cgt aag tat aga gaa 627
Ser Val Lys Asn His Trp His Val Val Met Ala Arg Lys Tyr Arg Glu
          190                      195                      200

cac tct tct gct tac cgt agg aga aag ctt atg agt aat aat cca ctt 675
His Ser Ser Ala Tyr Arg Arg Arg Lys Leu Met Ser Asn Asn Pro Leu
          205                      210                      215

aaa cct cac ctc acc aat aat cat cat cct aac cct aac cct aat tac 723
Lys Pro His Leu Thr Asn Asn His His Pro Asn Pro Asn Pro Asn Tyr
          220                      225                      230

cac tct ttt atc tcc act aat cat tac ttc gct cag cct ttc ccc gag 771
His Ser Phe Ile Ser Thr Asn His Tyr Phe Ala Gln Pro Phe Pro Glu
          235                      240                      245

ttt aat ttg act cat cac ctg gtt aat aat gcc cct atc acg agt gac 819
Phe Asn Leu Thr His His Leu Val Asn Asn Ala Pro Ile Thr Ser Asp
          250                      255                      260                      265

cat aac cag ctt gtg ttg cct ttc cat tgc ttt caa ggt tat gag aac 867
His Asn Gln Leu Val Leu Pro Phe His Cys Phe Gln Gly Tyr Glu Asn
          270                      275                      280

aat gaa cct ccg atg gtt gtg agt atg ttt ggc aac caa atg atg gtc 915
Asn Glu Pro Pro Met Val Val Ser Met Phe Gly Asn Gln Met Met Val
          285                      290                      295

ggc gat aac gtt ggt gcc acg tca gac gcg tta tgc aat att ccg cac 963
Gly Asp Asn Val Gly Ala Thr Ser Asp Ala Leu Cys Asn Ile Pro His
          300                      305                      310

att gac cct agt aac caa gag aaa ccg gag cca aat gat gca atg cat 1011
Ile Asp Pro Ser Asn Gln Glu Lys Pro Glu Pro Asn Asp Ala Met His
          315                      320                      325

tgg atc gga atg gac gcg gta gat gag gag gtg ttc gaa aag gct aag 1059
Trp Ile Gly Met Asp Ala Val Asp Glu Glu Val Phe Glu Lys Ala Lys
          330                      335                      340                      345

cag caa cca cat ttt ttc gat ttt ctt ggc ttg ggg acg gcg tga 1104
Gln Gln Pro His Phe Phe Asp Phe Leu Gly Leu Gly Thr Ala *
          350                      355

atgttgaaca aattggtgtt aatcagataa cgacagtggc 1144

<210> 32
<211> 359
<212> PRT
<213> Arabidopsis thaliana

<220>
<221> DOMAIN
<222> (90)...(210)
<223> Conserved domain

<400> 32

```

```

Met Phe Ile Thr Glu Lys Gln Val Trp Met Asp Glu Ile Val Ala Arg
 1          5          10          15
Arg Ala Ser Ser Trp Asp Phe Pro Phe Asn Asp Ile Asn Ile His
 20          25          30
Gln His His His Arg His Cys Asn Thr Ser His Glu Phe Glu Ile Leu
 35          40          45
Lys Ser Pro Leu Gly Asp Val Ala Val His Glu Glu Glu Ser Asn Asn
 50          55          60
Asn Asn Pro Asn Phe Ser Asn Ser Glu Ser Gly Lys Lys Glu Thr Thr
 65          70          75          80
Asp Ser Gly Gln Ser Trp Ser Ser Ser Ser Ser Lys Pro Ser Val Leu
 85          90          95
Gly Arg Gly His Trp Arg Pro Ala Glu Asp Val Lys Leu Lys Glu Leu
 100          105          110
Val Ser Ile Tyr Gly Pro Gln Asn Trp Asn Leu Ile Ala Glu Lys Leu
 115          120          125
Gln Gly Arg Ser Gly Lys Ser Cys Arg Leu Arg Trp Phe Asn Gln Leu
 130          135          140
Asp Pro Arg Ile Asn Arg Arg Ala Phe Thr Glu Glu Glu Glu Arg
 145          150          155          160
Leu Met Gln Ala His Arg Leu Tyr Gly Asn Lys Trp Ala Met Ile Ala
 165          170          175
Arg Leu Phe Pro Gly Arg Thr Asp Asn Ser Val Lys Asn His Trp His
 180          185          190
Val Val Met Ala Arg Lys Tyr Arg Glu His Ser Ser Ala Tyr Arg Arg
 195          200          205
Arg Lys Leu Met Ser Asn Asn Pro Leu Lys Pro His Leu Thr Asn Asn
 210          215          220
His His Pro Asn Pro Asn Pro Asn Tyr His Ser Phe Ile Ser Thr Asn
 225          230          235          240
His Tyr Phe Ala Gln Pro Phe Pro Glu Phe Asn Leu Thr His His Leu
 245          250          255
Val Asn Asn Ala Pro Ile Thr Ser Asp His Asn Gln Leu Val Leu Pro
 260          265          270
Phe His Cys Phe Gln Gly Tyr Glu Asn Asn Glu Pro Pro Met Val Val
 275          280          285
Ser Met Phe Gly Asn Gln Met Met Val Gly Asp Asn Val Gly Ala Thr
 290          295          300
Ser Asp Ala Leu Cys Asn Ile Pro His Ile Asp Pro Ser Asn Gln Glu
 305          310          315          320
Lys Pro Glu Pro Asn Asp Ala Met His Trp Ile Gly Met Asp Ala Val
 325          330          335
Asp Glu Glu Val Phe Glu Lys Ala Lys Gln Gln Pro His Phe Phe Asp
 340          345          350
Phe Leu Gly Leu Gly Thr Ala
 355

```

&lt;210&gt; 33

&lt;211&gt; 1009

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (177) ... (866)

&lt;400&gt; 33

cattccgaaa agtacttggt tattggtggt attcatcgct gctcaagggt gcaaaccaag 60  
 gagaacaaga agagtcagg gccaatgagc ctgagggtat caaatcagtg attgatgaca 120  
 tctacatttt tgaagacgac cctgtggttg ataggttcaa taggtttggc acggag atg 179

Met

1

tct gcc atg gtc tgc gcc ttg aca cag gtg gtt tct gct cgc tct cag 227  
 Ser Ala Met Val Ser Ala Leu Thr Gln Val Val Ser Ala Arg Ser Gln  
                   5                                  10                                  15

act gag gct gaa ggt gct cac tct cag act gag gct gaa ggt gct cac 275  
 Thr Glu Ala Glu Gly Ala His Ser Gln Thr Glu Ala Glu Gly Ala His  
                   20                                  25                                  30

tct tct tcc tct tgc gct gga cat aaa aga gga tgg ctt gga atc gat 323  
 Ser Ser Ser Ser Ser Ala Gly His Lys Arg Gly Trp Leu Gly Ile Asp  
                   35                                  40                                  45

tct gct cct att ccc tca tca ttt gct cgt gta gac tct tca cat aat 371  
 Ser Ala Pro Ile Pro Ser Ser Phe Ala Arg Val Asp Ser Ser His Asn  
                   50                                  55                                  60                                  65

ccg atc gaa gaa tcc atg agc aag gca ttt cca gag gaa gca agg gag 419  
 Pro Ile Glu Glu Ser Met Ser Lys Ala Phe Pro Glu Glu Ala Arg Glu  
                                   70                                  75                                  80

aaa aaa agg agg tac aga gga gta agg cag aga cca tgg ggc aaa tgg 467  
 Lys Lys Arg Arg Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Trp  
                   85                                  90                                  95

gca gct gag ata cgt gat cca cat aga gcc gct agg gtt tgg ctc ggg 515  
 Ala Ala Glu Ile Arg Asp Pro His Arg Ala Ala Arg Val Trp Leu Gly  
                   100                                  105                                  110

acg ttt gat aca gcg gag gcc gcg gct aga gcc tac gac gag gct gca 563  
 Thr Phe Asp Thr Ala Glu Ala Ala Ala Arg Ala Tyr Asp Glu Ala Ala  
                   115                                  120                                  125

ctc cgg ttc cgt gga aat aaa gca aag cta aat ttc cca gag gat gta 611  
 Leu Arg Phe Arg Gly Asn Lys Ala Lys Leu Asn Phe Pro Glu Asp Val  
                   130                                  135                                  140                                  145

agg att ctt cct cct ccc cct cct ctt ctt cgt tca cca gct gac acg 659  
 Arg Ile Leu Pro Pro Pro Pro Pro Leu Leu Arg Ser Pro Ala Asp Thr  
                                   150                                  155                                  160

gtg gcg aat aaa gca gaa gag gat ctg ata aat tat tgg agt tat aca 707  
 Val Ala Asn Lys Ala Glu Glu Asp Leu Ile Asn Tyr Trp Ser Tyr Thr  
                   165                                  170                                  175

aag ttg ttg caa agt tca ggc caa cgg tca ttt ctc gag cga gga caa 755  
 Lys Leu Leu Gln Ser Ser Gly Gln Arg Ser Phe Leu Glu Arg Gly Gln  
                   180                                  185                                  190

gaa gag agt agt aac ata ttt gaa cat tca cca atg gaa caa cct ctg 803  
 Glu Glu Ser Ser Asn Ile Phe Glu His Ser Pro Met Glu Gln Pro Leu

195                      200                      205

cct cct tca agt tct ggt cca agt tcc tct aat ttt cct gca cct tct 851  
 Pro Pro Ser Ser Ser Gly Pro Ser Ser Ser Asn Phe Pro Ala Pro Ser  
 210                      215                      220                      225

cta cct aat aca tag ttccactttt attaaagctc tacaaataca attaaataca 906  
 Leu Pro Asn Thr \*

tagctaaatg aaaatgattt tcttgtctgt ataccttctt aagtgcataa caatatattg 966  
 tactctttgt ttctcttaac taaaaaaaaa aaaaaaaaaa aaa 1009

<210> 34  
 <211> 229  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 34

Met	Ser	Ala	Met	Val	Ser	Ala	Leu	Thr	Gln	Val	Val	Ser	Ala	Arg	Ser
1			5						10					15	
Gln	Thr	Glu	Ala	Glu	Gly	Ala	His	Ser	Gln	Thr	Glu	Ala	Glu	Gly	Ala
		20					25						30		
His	Ser	Ser	Ser	Ser	Ser	Ala	Gly	His	Lys	Arg	Gly	Trp	Leu	Gly	Ile
		35					40					45			
Asp	Ser	Ala	Pro	Ile	Pro	Ser	Ser	Phe	Ala	Arg	Val	Asp	Ser	Ser	His
	50					55					60				
Asn	Pro	Ile	Glu	Glu	Ser	Met	Ser	Lys	Ala	Phe	Pro	Glu	Glu	Ala	Arg
65				70					75					80	
Glu	Lys	Lys	Arg	Arg	Tyr	Arg	Gly	Val	Arg	Gln	Arg	Pro	Trp	Gly	Lys
		85						90						95	
Trp	Ala	Ala	Glu	Ile	Arg	Asp	Pro	His	Arg	Ala	Ala	Arg	Val	Trp	Leu
		100						105					110		
Gly	Thr	Phe	Asp	Thr	Ala	Glu	Ala	Ala	Arg	Ala	Tyr	Asp	Glu	Ala	
		115					120					125			
Ala	Leu	Arg	Phe	Arg	Gly	Asn	Lys	Ala	Lys	Leu	Asn	Phe	Pro	Glu	Asp
		130				135					140				
Val	Arg	Ile	Leu	Pro	Pro	Pro	Pro	Pro	Leu	Leu	Arg	Ser	Pro	Ala	Asp
145				150					155					160	
Thr	Val	Ala	Asn	Lys	Ala	Glu	Glu	Asp	Leu	Ile	Asn	Tyr	Trp	Ser	Tyr
			165						170					175	
Thr	Lys	Leu	Leu	Gln	Ser	Ser	Gly	Gln	Arg	Ser	Phe	Leu	Glu	Arg	Gly
		180						185					190		
Gln	Glu	Glu	Ser	Ser	Asn	Ile	Phe	Glu	His	Ser	Pro	Met	Glu	Gln	Pro
		195					200					205			
Leu	Pro	Pro	Ser	Ser	Ser	Gly	Pro	Ser	Ser	Ser	Asn	Phe	Pro	Ala	Pro
	210					215					220				
Ser	Leu	Pro	Asn	Thr											
225															

<210> 35  
 <211> 1997  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>



&lt;221&gt; CDS

&lt;222&gt; (69)...(1907)

&lt;400&gt; 35

```

aacttcaaca tacacataat ctctcactta aaaatatctc tctctctctc tctacaaaat 60
caattcca atg ttg gtg gga aag ata agt gga tat gaa gat aat act cgc 110
  \  Met Leu Val Gly Lys Ile Ser Gly Tyr Glu Asp Asn Thr Arg
    1          5          10

tct ttg gag cga gaa aca tct gaa atc act tct ctt ctc agc caa ttt 158
Ser Leu Glu Arg Glu Thr Ser Glu Ile Thr Ser Leu Leu Ser Gln Phe
 15          20          25          30

ccg ggg aat act aat gtc ctt gtt gtt gac acc aat ttc acc act cta 206
Pro Gly Asn Thr Asn Val Leu Val Val Asp Thr Asn Phe Thr Thr Leu
          35          40          45

ctc aac atg aaa caa atc atg aaa caa tac gct tat caa gtg tct att 254
Leu Asn Met Lys Gln Ile Met Lys Gln Tyr Ala Tyr Gln Val Ser Ile
          50          55          60

gag aca gat gca gaa aaa gct ctt gcg ttt ttg aca agc tgc aaa cat 302
Glu Thr Asp Ala Glu Lys Ala Leu Ala Phe Leu Thr Ser Cys Lys His
          65          70          75

gaa atc aat att gtg att tgg gat ttt cat atg cct gga att gat gga 350
Glu Ile Asn Ile Val Ile Trp Asp Phe His Met Pro Gly Ile Asp Gly
          80          85          90

ctt caa gct ctc aag agc att act tca aag ttg gat tta cct gta gtg 398
Leu Gln Ala Leu Lys Ser Ile Thr Ser Lys Leu Asp Leu Pro Val Val
          95          100          105          110

att atg tct gat gat aat caa acg gaa tct gtg atg aaa gca aca ttt 446
Ile Met Ser Asp Asp Asn Gln Thr Glu Ser Val Met Lys Ala Thr Phe
          115          120          125

tac ggt gct tgt gac tat gtt gtg aaa ccg gtt aaa gaa gag gta atg 494
Tyr Gly Ala Cys Asp Tyr Val Val Lys Pro Val Lys Glu Glu Val Met
          130          135          140

gcc aat ata tgg caa cac att gta cgg aag agg ctg atc ttt aaa ccg 542
Ala Asn Ile Trp Gln His Ile Val Arg Lys Arg Leu Ile Phe Lys Pro
          145          150          155

gat gtt gct cca ccg gtt caa tca gat ccg gct cgc tct gac cgt tta 590
Asp Val Ala Pro Pro Val Gln Ser Asp Pro Ala Arg Ser Asp Arg Leu
          160          165          170

gac caa gtc aaa gct gat ttc aag atc gta gaa gat gaa cca ata atc 638
Asp Gln Val Lys Ala Asp Phe Lys Ile Val Glu Asp Glu Pro Ile Ile
          175          180          185          190

aat gag aca ccg ctg atc aca tgg acc gaa gaa att caa ccg gtt cag 686
Asn Glu Thr Pro Leu Ile Thr Trp Thr Glu Glu Ile Gln Pro Val Gln
          195          200          205

```

tca gat ctg gtt caa gcc aac aag ttc gac caa gtg aat ggc tat tcc	734
Ser Asp Leu Val Gln Ala Asn Lys Phe Asp Gln Val Asn Gly Tyr Ser	
210 215 220	
cca atc atg aac caa gat aac atg ttc aac aaa gca cca cct aaa ccg	782
Pro Ile Met Asn Gln Asp Asn Met Phe Asn Lys Ala Pro Pro Lys Pro	
225 230 235	
cga atg acg tgg aca gaa gtt att caa ccg gtt caa tca aat ctg gtt	830
Arg Met Thr Trp Thr Glu Val Ile Gln Pro Val Gln Ser Asn Leu Val	
240 245 250	
caa aca aaa gag ttc ggc caa ctc aat gac tat tcc caa atc atg aac	878
Gln Thr Lys Glu Phe Gly Gln Leu Asn Asp Tyr Ser Gln Ile Met Asn	
255 260 265 270	
caa gat agc atg tac aac aaa gca gca acc aaa cca caa ttg acg tgg	926
Gln Asp Ser Met Tyr Asn Lys Ala Ala Thr Lys Pro Gln Leu Thr Trp	
275 280 285	
acc gaa gaa att caa ccg gtt caa tca ggt ctg gtt caa gcc aac gag	974
Thr Glu Glu Ile Gln Pro Val Gln Ser Gly Leu Val Gln Ala Asn Glu	
290 295 300	
ttc agc aaa gtg aat gga tat tcc caa agc atg aac caa gat agc atg	1022
Phe Ser Lys Val Asn Gly Tyr Ser Gln Ser Met Asn Gln Asp Ser Met	
305 310 315	
ttc aac aaa tca gca acc aac ccg cga ttg aca tgg aac gaa tta ctt	1070
Phe Asn Lys Ser Ala Thr Asn Pro Arg Leu Thr Trp Asn Glu Leu Leu	
320 325 330	
caa ccg gtt caa tca gat ctg gtt caa tcc aat gag ttt agc caa ttc	1118
Gln Pro Val Gln Ser Asp Leu Val Gln Ser Asn Glu Phe Ser Gln Phe	
335 340 345 350	
agt gac tat tct caa atc atg aac gaa gat aac atg ttc aac aaa gca	1166
Ser Asp Tyr Ser Gln Ile Met Asn Glu Asp Asn Met Phe Asn Lys Ala	
355 360 365	
gca aag aaa ccg cgg atg aca tgg agt gaa gta ttt caa ccg gtt caa	1214
Ala Lys Lys Pro Arg Met Thr Trp Ser Glu Val Phe Gln Pro Val Gln	
370 375 380	
tca cat ctg gtt ccg act gac ggt tta gac cga gac cac ttt gat tcc	1262
Ser His Leu Val Pro Thr Asp Gly Leu Asp Arg Asp His Phe Asp Ser	
385 390 395	
ata acc ata aac gga ggt aac ggc ata caa aac atg gaa aag aaa caa	1310
Ile Thr Ile Asn Gly Gly Asn Gly Ile Gln Asn Met Glu Lys Lys Gln	
400 405 410	
gga aaa aaa cca cgg aag ccg cgg atg acg tgg acc gaa gag ctt cac	1358
Gly Lys Lys Pro Arg Lys Pro Arg Met Thr Trp Thr Glu Glu Leu His	
415 420 425 430	
caa aaa ttt ctg gaa gcc atc gaa ata att ggt ggt atc gaa aaa gct	1406

Gln Lys Phe Leu Glu Ala Ile Glu Ile Ile Gly Gly Ile Glu Lys Ala  
 435 440 445

aac cca aag gta ctt gtc gaa tgc ttg caa gaa atg agg ata gaa gga 1454  
 Asn Pro Lys Val Leu Val Glu Cys Leu Gln Glu Met Arg Ile Glu Gly  
 450 455 460

att act aga agc aat gtg gca agt cat ctt cag aaa cac cgt atc aat 1502  
 Ile Thr Arg Ser Asn Val Ala Ser His Leu Gln Lys His Arg Ile Asn  
 465 470 475

ctt gaa gaa aac caa att cct caa caa aca caa ggg aat ggt tgg gcc 1550  
 Leu Glu Glu Asn Gln Ile Pro Gln Gln Thr Gln Gly Asn Gly Trp Ala  
 480 485 490

act gcg tat ggt aca cta gct ccc tct ctc caa ggt tca gac aat gtc 1598  
 Thr Ala Tyr Gly Thr Leu Ala Pro Ser Leu Gln Gly Ser Asp Asn Val  
 495 500 505 510

aac aca aca ata cca tgc tac ctt atg aat ggt cca gcc act ttg aac 1646  
 Asn Thr Thr Ile Pro Ser Tyr Leu Met Asn Gly Pro Ala Thr Leu Asn  
 515 520 525

caa atc cag cag aat caa tat caa aat ggt ttc ttg aca atg aac aac 1694  
 Gln Ile Gln Gln Asn Gln Tyr Gln Asn Gly Phe Leu Thr Met Asn Asn  
 530 535 540

aac cag atc ata acc aat cct ccg cct cct ttg ccc tat ttg gac cat 1742  
 Asn Gln Ile Ile Thr Asn Pro Pro Pro Pro Leu Pro Tyr Leu Asp His  
 545 550 555

cat cac caa cag caa cat cag tct tct cct caa ttt aat tac ctg atg 1790  
 His His Gln Gln Gln His Gln Ser Ser Pro Gln Phe Asn Tyr Leu Met  
 560 565 570

aac aat gaa gaa ctt ctt caa gcc tct ggc ctc tct gcg aca gat ctt 1838  
 Asn Asn Glu Glu Leu Leu Gln Ala Ser Gly Leu Ser Ala Thr Asp Leu  
 575 580 585 590

gaa ctc act tat cca agt tta cca tat gat cca caa gag tat cta atc 1886  
 Glu Leu Thr Tyr Pro Ser Leu Pro Tyr Asp Pro Gln Glu Tyr Leu Ile  
 595 600 605

aat ggc tac aat tat aat tag tcatatagcc cttctcttta cttaaggcag 1937  
 Asn Gly Tyr Asn Tyr Asn \*  
 610

tctatgtatg acaaataata tgcgacttcc cttgtgagtc acaatattgt ttcattattc 1997

<210> 36

<211> 612

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

&lt;222&gt; (425)...(478)

&lt;223&gt; Conserved domain

&lt;400&gt; 36

```

Met Leu Val Gly Lys Ile Ser Gly Tyr Glu Asp Asn Thr Arg Ser Leu
 1           5           10           15
Glu Arg Glu Thr Ser Glu Ile Thr Ser Leu Leu Ser Gln Phe Pro Gly
          20           25           30
Asn Thr Asn Val Leu Val Val Asp Thr Asn Phe Thr Thr Leu Leu Asn
          35           40           45
Met Lys Gln Ile Met Lys Gln Tyr Ala Tyr Gln Val Ser Ile Glu Thr
          50           55           60
Asp Ala Glu Lys Ala Leu Ala Phe Leu Thr Ser Cys Lys His Glu Ile
65           70           75           80
Asn Ile Val Ile Trp Asp Phe His Met Pro Gly Ile Asp Gly Leu Gln
          85           90           95
Ala Leu Lys Ser Ile Thr Ser Lys Leu Asp Leu Pro Val Val Ile Met
          100          105          110
Ser Asp Asp Asn Gln Thr Glu Ser Val Met Lys Ala Thr Phe Tyr Gly
          115          120          125
Ala Cys Asp Tyr Val Val Lys Pro Val Lys Glu Glu Val Met Ala Asn
          130          135          140
Ile Trp Gln His Ile Val Arg Lys Arg Leu Ile Phe Lys Pro Asp Val
145          150          155          160
Ala Pro Pro Val Gln Ser Asp Pro Ala Arg Ser Asp Arg Leu Asp Gln
          165          170          175
Val Lys Ala Asp Phe Lys Ile Val Glu Asp Glu Pro Ile Ile Asn Glu
          180          185          190
Thr Pro Leu Ile Thr Trp Thr Glu Glu Ile Gln Pro Val Gln Ser Asp
          195          200          205
Leu Val Gln Ala Asn Lys Phe Asp Gln Val Asn Gly Tyr Ser Pro Ile
          210          215          220
Met Asn Gln Asp Asn Met Phe Asn Lys Ala Pro Pro Lys Pro Arg Met
225          230          235          240
Thr Trp Thr Glu Val Ile Gln Pro Val Gln Ser Asn Leu Val Gln Thr
          245          250          255
Lys Glu Phe Gly Gln Leu Asn Asp Tyr Ser Gln Ile Met Asn Gln Asp
          260          265          270
Ser Met Tyr Asn Lys Ala Ala Thr Lys Pro Gln Leu Thr Trp Thr Glu
          275          280          285
Glu Ile Gln Pro Val Gln Ser Gly Leu Val Gln Ala Asn Glu Phe Ser
          290          295          300
Lys Val Asn Gly Tyr Ser Gln Ser Met Asn Gln Asp Ser Met Phe Asn
305          310          315          320
Lys Ser Ala Thr Asn Pro Arg Leu Thr Trp Asn Glu Leu Leu Gln Pro
          325          330          335
Val Gln Ser Asp Leu Val Gln Ser Asn Glu Phe Ser Gln Phe Ser Asp
          340          345          350
Tyr Ser Gln Ile Met Asn Glu Asp Asn Met Phe Asn Lys Ala Ala Lys
          355          360          365
Lys Pro Arg Met Thr Trp Ser Glu Val Phe Gln Pro Val Gln Ser His
          370          375          380
Leu Val Pro Thr Asp Gly Leu Asp Arg Asp His Phe Asp Ser Ile Thr
385          390          395          400
Ile Asn Gly Gly Asn Gly Ile Gln Asn Met Glu Lys Lys Gln Gly Lys
          405          410          415
Lys Pro Arg Lys Pro Arg Met Thr Trp Thr Glu Glu Leu His Gln Lys

```

```

      420      425      430
Phe Leu Glu Ala Ile Glu Ile Ile Gly Gly Ile Glu Lys Ala Asn Pro
      435      440      445
Lys Val Leu Val Glu Cys Leu Gln Glu Met Arg Ile Glu Gly Ile Thr
      450      455      460
Arg Ser Asn Val Ala Ser His Leu Gln Lys His Arg Ile Asn Leu Glu
465      470      475      480
Glu Asn Gln Ile Pro Gln Gln Thr Gln Gly Asn Gly Trp Ala Thr Ala
      485      490      495
Tyr Gly Thr Leu Ala Pro Ser Leu Gln Gly Ser Asp Asn Val Asn Thr
      500      505      510
Thr Ile Pro Ser Tyr Leu Met Asn Gly Pro Ala Thr Leu Asn Gln Ile
      515      520      525
Gln Gln Asn Gln Tyr Gln Asn Gly Phe Leu Thr Met Asn Asn Asn Gln
      530      535      540
Ile Ile Thr Asn Pro Pro Pro Pro Leu Pro Tyr Leu Asp His His His
545      550      555      560
Gln Gln Gln His Gln Ser Ser Pro Gln Phe Asn Tyr Leu Met Asn Asn
      565      570      575
Glu Glu Leu Leu Gln Ala Ser Gly Leu Ser Ala Thr Asp Leu Glu Leu
      580      585      590
Thr Tyr Pro Ser Leu Pro Tyr Asp Pro Gln Glu Tyr Leu Ile Asn Gly
      595      600      605
Tyr Asn Tyr Asn
      610

```

<210> 37  
 <211> 759  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (16)...(477)

```

<400> 37
gtttgtcttt tatca atg gaa aga gaa caa gaa gag tct acg atg aga aag 51
      Met Glu Arg Glu Gln Glu Glu Ser Thr Met Arg Lys
      1      5      10

aga agg cag cca cct caa gaa gaa gtg cct aac cac gtg gct aca agg 99
Arg Arg Gln Pro Pro Gln Glu Glu Val Pro Asn His Val Ala Thr Arg
      15      20      25

aag ccg tac aga ggg ata cgg agg agg aag tgg ggc aag tgg gtg gct 147
Lys Pro Tyr Arg Gly Ile Arg Arg Arg Lys Trp Gly Lys Trp Val Ala
      30      35      40

gag att cgt gag cct aac aaa cgc tca cgg ctt tgg ctt ggc tct tac 195
Glu Ile Arg Glu Pro Asn Lys Arg Ser Arg Leu Trp Leu Gly Ser Tyr
      45      50      55      60

aca acc gat atc gcc gcc gct aga gcc tac gac gtg gcc gtc ttc tac 243
Thr Thr Asp Ile Ala Ala Arg Ala Tyr Asp Val Ala Val Phe Tyr
      65      70      75

```

ctc cgt ggc ccc tcc gca cgt ctc aac ttc cct gat ctt ctc ttg caa 291  
 Leu Arg Gly Pro Ser Ala Arg Leu Asn Phe Pro Asp Leu Leu Leu Gln  
                     80                    85                    90

gaa gag gac cat ctc tca gcc gcc acc acc gct gac atg ccc gca gct 339  
 Glu Glu Asp His Leu Ser Ala Ala Thr Thr Ala Asp Met Pro Ala Ala  
                     95                    100                    105

ctt ata agg gaa aaa gcg gcg gag gtc ggc gcc aga gtc gac gct ctt 387  
 Leu Ile Arg Glu Lys Ala Ala Glu Val Gly Ala Arg Val Asp Ala Leu  
                     110                    115                    120

cta gct tct gcc gct cct tcg atg gct cac tcc act ccg ccg gta ata 435  
 Leu Ala Ser Ala Ala Pro Ser Met Ala His Ser Thr Pro Pro Val Ile  
                     125                    130                    135                    140

aaa ccc gac ttg aat caa ata ccc gaa tcc gga gat ata tag 477  
 Lys Pro Asp Leu Asn Gln Ile Pro Glu Ser Gly Asp Ile \*  
                     145                    150

tcaatttata tacatgtagt ttgttttgtt tgattagaag attacattta catacaagat 537  
 acacatagat actggaaaat ataggtatgt atacattcat aaattatctt atgtatcaaa 597  
 gaattttata gattctgatt agctttttgt ttttgttttt gataagaact ctgattagtt 657  
 gtccggagac aaaaccggct aagagcaatc catgagaagc tagcgagtgt tttttagttc 717  
 aagttgtaat ataaatgcat attaattctt tagtaatttt gt 759

<210> 38  
 <211> 153  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (28) ... (95)  
 <223> Conserved domain

<400> 38  
 Met Glu Arg Glu Gln Glu Glu Ser Thr Met Arg Lys Arg Arg Gln Pro  
   1                    5                    10                    15  
 Pro Gln Glu Glu Val Pro Asn His Val Ala Thr Arg Lys Pro Tyr Arg  
                     20                    25                    30  
 Gly Ile Arg Arg Arg Lys Trp Gly Lys Trp Val Ala Glu Ile Arg Glu  
                     35                    40                    45  
 Pro Asn Lys Arg Ser Arg Leu Trp Leu Gly Ser Tyr Thr Thr Asp Ile  
                     50                    55                    60  
 Ala Ala Ala Arg Ala Tyr Asp Val Ala Val Phe Tyr Leu Arg Gly Pro  
   65                    70                    75                    80  
 Ser Ala Arg Leu Asn Phe Pro Asp Leu Leu Leu Gln Glu Glu Asp His  
                     85                    90                    95  
 Leu Ser Ala Ala Thr Thr Ala Asp Met Pro Ala Ala Leu Ile Arg Glu  
                     100                    105                    110  
 Lys Ala Ala Glu Val Gly Ala Arg Val Asp Ala Leu Leu Ala Ser Ala  
                     115                    120                    125  
 Ala Pro Ser Met Ala His Ser Thr Pro Pro Val Ile Lys Pro Asp Leu  
                     130                    135                    140  
 Asn Gln Ile Pro Glu Ser Gly Asp Ile  
   145                    150

&lt;210&gt; 39

&lt;211&gt; 992

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (41)...(760)

&lt;400&gt; 39

```

ctctgctggg atcattggag tctagggttt tgttattgac atg cgt ggt gtg tca 55
                                         Met Arg Gly Val Ser
                                         1       5

gaa ttg gag gtg ggg aag agt aat ctt ccg gcg gag agt gag ctg gaa 103
Glu Leu Glu Val Gly Lys Ser Asn Leu Pro Ala Glu Ser Glu Leu Glu
              10              15              20

ttg gga tta ggg ctc agc ctc ggt ggt ggc gcg tgg aaa gag cgt ggg 151
Leu Gly Leu Gly Leu Ser Leu Gly Gly Gly Ala Trp Lys Glu Arg Gly
              25              30              35

agg att ctt act gct aag gat ttt cct tcc gtt ggg tct aaa cgc tct 199
Arg Ile Leu Thr Ala Lys Asp Phe Pro Ser Val Gly Ser Lys Arg Ser
              40              45              50

gct gaa tct tcc tct cac caa gga gct tct cct cct cgt tca agt caa 247
Ala Glu Ser Ser Ser His Gln Gly Ala Ser Pro Pro Arg Ser Ser Gln
              55              60              65

gtg gta gga tgg cca cca att ggg tta cac agg atg aac agt ttg gtt 295
Val Val Gly Trp Pro Pro Ile Gly Leu His Arg Met Asn Ser Leu Val
              70              75              80              85

aat aac caa gct atg aag gca gca aga gcg gaa gaa gga gac ggg gag 343
Asn Asn Gln Ala Met Lys Ala Ala Arg Ala Glu Glu Gly Asp Gly Glu
              90              95              100

aag aaa gtt gtg aag aat gat gag ctc aaa gat gtg tca atg aag gtg 391
Lys Lys Val Val Lys Asn Asp Glu Leu Lys Asp Val Ser Met Lys Val
              105              110              115

aat ccg aaa gtt cag ggc tta ggg ttt gtt aag gtg aat atg gat gga 439
Asn Pro Lys Val Gln Gly Leu Gly Phe Val Lys Val Asn Met Asp Gly
              120              125              130

gtt ggt ata ggc aga aaa gtg gat atg aga gct cat tcg tct tac gaa 487
Val Gly Ile Gly Arg Lys Val Asp Met Arg Ala His Ser Ser Tyr Glu
              135              140              145

aac ttg gct cag acg ctt gag gaa atg ttc ttt gga atg aca ggt act 535
Asn Leu Ala Gln Thr Leu Glu Glu Met Phe Phe Gly Met Thr Gly Thr
              150              155              160              165

act tgt cga gaa aag gtt aaa cct tta agg ctt tta gat gga tca tca 583

```

Thr Cys Arg Glu Lys Val Lys Pro Leu Arg Leu Leu Asp Gly Ser Ser  
 170 175 180

gac ttt gta ctc act tat gaa gat aag gaa ggg gat tgg atg ctt gtt 631  
 Asp Phe Val Leu Thr Tyr Glu Asp Lys Glu Gly Asp Trp Met Leu Val  
 185 190 195

gga gat gtt cca tgg aga atg ttt atc aac tcg gtg aaa agg ctt cgg 679  
 Gly Asp Val Pro Trp Arg Met Phe Ile Asn Ser Val Lys Arg Leu Arg  
 200 205 210

atc atg gga acc tca gaa gct agt gga cta gct cca aga cgt caa gag 727  
 Ile Met Gly Thr Ser Glu Ala Ser Gly Leu Ala Pro Arg Arg Gln Glu  
 215 220 225

cag aag gat aga caa aga aac aac cct gtt tag cttcccttcc aaagctggca 780  
 Gln Lys Asp Arg Gln Arg Asn Asn Pro Val \*  
 230 235

ttgtttatgt attgtttgag gtttgcaatt tactcgatac tttttgaaga aagtattttg 840  
 gagaatatgg ataaaagcat gcagaagctt agatatgatt tgaatccggt ttccggatat 900  
 gggttttgctt aggtcattca attcgtagtt ttccagtttg tttcttcttt ggctgtgtac 960  
 caattatcta tgttctgtga gagaaagctc tt 992

<210> 40  
 <211> 239  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (20)...(28)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (71)...(82)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (126)...(142)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (187)...(224)  
 <223> Conserved domain

<400> 40  
 Met Arg Gly Val Ser Glu Leu Glu Val Gly Lys Ser Asn Leu Pro Ala  
 1 5 10 15  
 Glu Ser Glu Leu Glu Leu Gly Leu Gly Leu Ser Leu Gly Gly Gly Ala  
 20 25 30  
 Trp Lys Glu Arg Gly Arg Ile Leu Thr Ala Lys Asp Phe Pro Ser Val  
 35 40 45  
 Gly Ser Lys Arg Ser Ala Glu Ser Ser Ser His Gln Gly Ala Ser Pro  
 50 55 60  
 Pro Arg Ser Ser Gln Val Val Gly Trp Pro Pro Ile Gly Leu His Arg  
 65 70 75 80



Met Asn Ser Leu Val Asn Asn Gln Ala Met Lys Ala Ala Arg Ala Glu  
                             85                            90                            95  
 Glu Gly Asp Gly Glu Lys Lys Val Val Lys Asn Asp Glu Leu Lys Asp  
                             100                            105                            110  
 Val Ser Met Lys Val Asn Pro Lys Val Gln Gly Leu Gly Phe Val Lys  
                             115                            120                            125  
 Val Asn Met Asp Gly Val Gly Ile Gly Arg Lys Val Asp Met Arg Ala  
                             130                            135                            140  
 His Ser Ser Tyr Glu Asn Leu Ala Gln Thr Leu Glu Glu Met Phe Phe  
                             145                            150                            155                            160  
 Gly Met Thr Gly Thr Thr Cys Arg Glu Lys Val Lys Pro Leu Arg Leu  
                             165                            170                            175  
 Leu Asp Gly Ser Ser Asp Phe Val Leu Thr Tyr Glu Asp Lys Glu Gly  
                             180                            185                            190  
 Asp Trp Met Leu Val Gly Asp Val Pro Trp Arg Met Phe Ile Asn Ser  
                             195                            200                            205  
 Val Lys Arg Leu Arg Ile Met Gly Thr Ser Glu Ala Ser Gly Leu Ala  
                             210                            215                            220  
 Pro Arg Arg Gln Glu Gln Lys Asp Arg Gln Arg Asn Asn Pro Val  
                             225                            230                            235

&lt;210&gt; 41

&lt;211&gt; 228

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(228)

&lt;400&gt; 41

atg gat aac cat cgc agg act aag caa ccc aag acc aac tcc atc gtt 48  
 Met Asp Asn His Arg Arg Thr Lys Gln Pro Lys Thr Asn Ser Ile Val  
   1                            5                            10                            15

act tct tct tct gaa gaa gtg agt agt ctt gag tgg gaa gtt gtg aac 96  
 Thr Ser Ser Ser Glu Glu Val Ser Ser Leu Glu Trp Glu Val Val Asn  
                             20                            25                            30

atg agt caa gaa gaa gaa gat ttg gtc tct cga atg cat aag ctt gtc 144  
 Met Ser Gln Glu Glu Glu Asp Leu Val Ser Arg Met His Lys Leu Val  
                             35                            40                            45

ggt gac agg tgg gaa ctg ata gct ggg agg atc cca gga aga acc gct 192  
 Gly Asp Arg Trp Glu Leu Ile Ala Gly Arg Ile Pro Gly Arg Thr Ala  
                             50                            55                            60

gga gaa att gag agg ttt tgg gtc atg aaa aat tga 228  
 Gly Glu Ile Glu Arg Phe Trp Val Met Lys Asn \*  
   65                            70                            75

&lt;210&gt; 42

&lt;211&gt; 75

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (27)...(63)  
 <223> Conserved domain

<400> 42  
 Met Asp Asn His Arg Arg Thr Lys Gln Pro Lys Thr Asn Ser Ile Val  
 1 5 10 15  
 Thr Ser Ser Ser Glu Glu Val Ser Ser Leu Glu Trp Glu Val Val Asn  
 20 25 30  
 Met Ser Gln Glu Glu Glu Asp Leu Val Ser Arg Met His Lys Leu Val  
 35 40 45  
 Gly Asp Arg Trp Glu Leu Ile Ala Gly Arg Ile Pro Gly Arg Thr Ala  
 50 55 60  
 Gly Glu Ile Glu Arg Phe Trp Val Met Lys Asn  
 65 70 75

<210> 43  
 <211> 1221  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (162)...(1013)

<400> 43  
 atttctcttc cacaaagagt cctaacttcg agttgaaaca aacaccattt ctcattctcta 60  
 tctcagaaag aacaaaccat ttcgtgttct ttctttctct attctcataa ggaaatataa 120  
 ttctgaaac tggtgagttc ttgtgaaagg aaataaaaaa c atg atg atg ggc aaa 176  
 Met Met Met Gly Lys  
 1 5  
 gaa gat cta ggt ttg agc cta agc tta ggg ttt tca caa aat cac aat 224  
 Glu Asp Leu Gly Leu Ser Leu Ser Leu Gly Phe Ser Gln Asn His Asn  
 10 15 20  
 cct ctt cag atg aat ctg aat cct aac tct tca tta tca aac aat ctc 272  
 Pro Leu Gln Met Asn Leu Asn Pro Asn Ser Ser Leu Ser Asn Asn Leu  
 25 30 35  
 cag aga ctc cca tgg aac caa aca ttc gat cct aca tca gat ctt cgc 320  
 Gln Arg Leu Pro Trp Asn Gln Thr Phe Asp Pro Thr Ser Asp Leu Arg  
 40 45 50  
 aag ata gac gtg aac agt ttt cca tca acg gtt aac tgc gag gaa gac 368  
 Lys Ile Asp Val Asn Ser Phe Pro Ser Thr Val Asn Cys Glu Glu Asp  
 55 60 65  
 aca gga gtt tcg tca cca aac agt acg atc tca agc acc att agc ggg 416  
 Thr Gly Val Ser Ser Pro Asn Ser Thr Ile Ser Ser Thr Ile Ser Gly  
 70 75 80 85  
 aag aga agt gag aga gaa gga atc tcc gga acc ggc gtt ggc tcc ggc 464  
 Lys Arg Ser Glu Arg Glu Gly Ile Ser Gly Thr Gly Val Gly Ser Gly

90	95	100	
gac gat cac gac gag atc act ccg gat cga ggg tac tca cgt gga acc			512
Asp Asp His Asp Glu Ile Thr Pro Asp Arg Gly Tyr Ser Arg Gly Thr			
105	110	115	
tca gat gaa gaa gaa gac ggg ggc gaa acg tcg agg aag aag ctc agg			560
Ser Asp Glu Glu Glu Asp Gly Gly Glu Thr Ser Arg Lys Lys Leu Arg			
120	125	130	
tta tca aaa gat cag tct gct ttt ctc gaa gag act ttc aaa gaa cac			608
Leu Ser Lys Asp Gln Ser Ala Phe Leu Glu Glu Thr Phe Lys Glu His			
135	140	145	
aac act ctc aat ccc aaa cag aag cta gct ttg gct aag aag ctg aac			656
Asn Thr Leu Asn Pro Lys Gln Lys Leu Ala Leu Ala Lys Lys Leu Asn			
150	155	160	165
ttg acg gca aga caa gtg gaa gtg tgg ttc caa aac aga aga gct aga			704
Leu Thr Ala Arg Gln Val Glu Val Trp Phe Gln Asn Arg Arg Ala Arg			
170	175	180	
acc aag tta aag caa acg gag gta gat tgc gaa tac ttg aaa cgg tgc			752
Thr Lys Leu Lys Gln Thr Glu Val Asp Cys Glu Tyr Leu Lys Arg Cys			
185	190	195	
gta gag aag cta acg gaa gag aac cgg aga ctt cag aaa gag gct atg			800
Val Glu Lys Leu Thr Glu Glu Asn Arg Arg Leu Gln Lys Glu Ala Met			
200	205	210	
gag ctt cga act ctc aag ctg tct cca caa ttc tac ggt cag atg act			848
Glu Leu Arg Thr Leu Lys Leu Ser Pro Gln Phe Tyr Gly Gln Met Thr			
215	220	225	
cca cca act aca ctc atc atg tgt cct tcg tgc gag cgt gta gct ggt			896
Pro Pro Thr Thr Leu Ile Met Cys Pro Ser Cys Glu Arg Val Ala Gly			
230	235	240	245
cca tca tca tcg aac cat cac cac aat cac agg ccg gtt tcg att aac			944
Pro Ser Ser Ser Asn His His Asn His Arg Pro Val Ser Ile Asn			
250	255	260	
ccg tgg att gct tgt gct ggt cag gtg gct cat ggg ctg aat ttt gaa			992
Pro Trp Ile Ala Cys Ala Gly Gln Val Ala His Gly Leu Asn Phe Glu			
265	270	275	
gcc ttg cgt cca cga tcg taa ttttttagtgg tgggggaagg gtgttttggg			1043
Ala Leu Arg Pro Arg Ser *			
280			
ttttttcatt atcggttatat agtctatctg tgtgggggtca ttgtaatttt ggatgattgg			1103
ccttctcatg aactagtcac atgtatgatg caaccttaaa aatatttcaa gtagcaaaac			1163
ttaattacaa acttgctata ttaaccacaaa attatgaaaa aaaaaaaaaa aaaaaaaaaa			1221

&lt;210&gt; 44

&lt;211&gt; 283

&lt;212&gt; PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (126)...(186)

<223> Conserved domain

<400> 44

```

Met Met Met Gly Lys Glu Asp Leu Gly Leu Ser Leu Ser Leu Gly Phe
 1          5          10          15
Ser Gln Asn His Asn Pro Leu Gln Met Asn Leu Asn Pro Asn Ser Ser
          20          25          30
Leu Ser Asn Asn Leu Gln Arg Leu Pro Trp Asn Gln Thr Phe Asp Pro
          35          40          45
Thr Ser Asp Leu Arg Lys Ile Asp Val Asn Ser Phe Pro Ser Thr Val
          50          55          60
Asn Cys Glu Glu Asp Thr Gly Val Ser Ser Pro Asn Ser Thr Ile Ser
65          70          75          80
Ser Thr Ile Ser Gly Lys Arg Ser Glu Arg Glu Gly Ile Ser Gly Thr
          85          90          95
Gly Val Gly Ser Gly Asp Asp His Asp Glu Ile Thr Pro Asp Arg Gly
          100          105          110
Tyr Ser Arg Gly Thr Ser Asp Glu Glu Glu Asp Gly Gly Glu Thr Ser
          115          120          125
Arg Lys Lys Leu Arg Leu Ser Lys Asp Gln Ser Ala Phe Leu Glu Glu
130          135          140
Thr Phe Lys Glu His Asn Thr Leu Asn Pro Lys Gln Lys Leu Ala Leu
145          150          155          160
Ala Lys Lys Leu Asn Leu Thr Ala Arg Gln Val Glu Val Trp Phe Gln
          165          170          175
Asn Arg Arg Ala Arg Thr Lys Leu Lys Gln Thr Glu Val Asp Cys Glu
          180          185          190
Tyr Leu Lys Arg Cys Val Glu Lys Leu Thr Glu Glu Asn Arg Arg Leu
          195          200          205
Gln Lys Glu Ala Met Glu Leu Arg Thr Leu Lys Leu Ser Pro Gln Phe
          210          215          220
Tyr Gly Gln Met Thr Pro Pro Thr Thr Leu Ile Met Cys Pro Ser Cys
225          230          235          240
Glu Arg Val Ala Gly Pro Ser Ser Ser Asn His His His Asn His Arg
          245          250          255
Pro Val Ser Ile Asn Pro Trp Ile Ala Cys Ala Gly Gln Val Ala His
          260          265          270
Gly Leu Asn Phe Glu Ala Leu Arg Pro Arg Ser
          275          280

```

<210> 45

<211> 696

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (77)...(496)

<400> 45

aatccataga tctcttatta aataacagtg ctgaccaagc tcttaccag caaaccaatc 60

```

tagaacacca aagtta atg gag agc tca aac agg agc agc aac aac caa tca 112
Met Glu Ser Ser Asn Arg Ser Ser Asn Asn Gln Ser
1 5 10

caa gat gac aag caa gct cgt ttc cgg gga gtt cga aga agg cct tgg 160
Gln Asp Asp Lys Gln Ala Arg Phe Arg Gly Val Arg Arg Arg Pro Trp
15 20 25

gga aag ttt gca gca gag att cga gac ccg tcg aga aac ggt gcc cgt 208
Gly Lys Phe Ala Ala Glu Ile Arg Asp Pro Ser Arg Asn Gly Ala Arg
30 35 40

ctt tgg ctc ggg aca ttt gag acc gct gag gag gca gca agg gct tat 256
Leu Trp Leu Gly Thr Phe Glu Thr Ala Glu Glu Ala Ala Arg Ala Tyr
45 50 55 60

gac cga gca gcc ttt aac ctt agg ggt cat ctc gct ata ctc aac ttc 304
Asp Arg Ala Ala Phe Asn Leu Arg Gly His Leu Ala Ile Leu Asn Phe
65 70 75

cct aat gag tat tat cca cgt atg gac gac tac tcg ctt cgc cct cct 352
Pro Asn Glu Tyr Tyr Pro Arg Met Asp Asp Tyr Ser Leu Arg Pro Pro
80 85 90

tat gct tct tct tct tcg tcg tcg tca tcg ggt tca act tct act aat 400
Tyr Ala Ser Ser Ser Ser Ser Ser Ser Ser Gly Ser Thr Ser Thr Asn
95 100 105

gtg agt cga caa aac caa aga gaa gtt ttc gag ttt gag tat ttg gac 448
Val Ser Arg Gln Asn Gln Arg Glu Val Phe Glu Phe Glu Tyr Leu Asp
110 115 120

gat aag gtt ctt gaa gaa ctt ctt gat tca gaa gaa agg aag aga taa 496
Asp Lys Val Leu Glu Glu Leu Leu Asp Ser Glu Glu Arg Lys Arg *
125 130 135

tcacgattag ttttgtttg atattttatg tggcactggt gtggctacct acgtgcatta 556
tgtgcatgta taggtcgctt gattagtact ttataacatg catgccacga ccataaattg 616
taagagaaga cgtactttgc gttttcatga aatatgaatg ttagatgggt tgagtacaaa 676
aaaaaaaaa aaaaaaaaaa 696

```

&lt;210&gt; 46

&lt;211&gt; 139

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (17)...(85)

&lt;223&gt; Conserved-domain

&lt;400&gt; 46

```

Met Glu Ser Ser Asn Arg Ser Ser Asn Asn Gln Ser Gln Asp Asp Lys
1 5 10 15
Gln Ala Arg Phe Arg Gly Val Arg Arg Arg Pro Trp Gly Lys Phe Ala
20 25 30
Ala Glu Ile Arg Asp Pro Ser Arg Asn Gly Ala Arg Leu Trp Leu Gly

```

```

      35              40              45
Thr Phe Glu Thr Ala Glu Glu Ala Ala Arg Ala Tyr Asp Arg Ala Ala
  50              55              60
Phe Asn Leu Arg Gly His Leu Ala Ile Leu Asn Phe Pro Asn Glu Tyr
  65              70              75              80
Tyr Pro Arg Met Asp Asp Tyr Ser Leu Arg Pro Pro Tyr Ala Ser Ser
      85              90              95
Ser Ser Ser Ser Ser Ser Gly Ser Thr Ser Thr Asn Val Ser Arg Gln
      100              105              110
Asn Gln Arg Glu Val Phe Glu Phe Glu Tyr Leu Asp Asp Lys Val Leu
      115              120              125
Glu Glu Leu Leu Asp Ser Glu Glu Arg Lys Arg
      130              135

```

&lt;210&gt; 47

&lt;211&gt; 1562

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (160)...(1335)

&lt;400&gt; 47

```

tctttctttc ttcctctttg tctctgttct ttgtttctct ctctctctct ctacagagtt 60
ttctttccct cgaagaaaaa gaatatTTTT aaatttaatt ttctctgcgt ttataagctt 120
taagtttcag aggaggagga tttagaagga gggttttgt atg tgt gtc tta aaa 174
                                Met Cys Val Leu Lys
                                1           5

```

```

gtg gca aat cag gaa gat aac gtt ggc aaa aaa gcc gag tct att aga 222
Val Ala Asn Gln Glu Asp Asn Val Gly Lys Lys Ala Glu Ser Ile Arg
      10              15              20

```

```

gac gat gat cat cgg acg tta tct gaa atc gat caa tgg ctt tac tta 270
Asp Asp Asp His Arg Thr Leu Ser Glu Ile Asp Gln Trp Leu Tyr Leu
      25              30              35

```

```

ttc gca gcc gaa gac gac cac cac cgt cat agc ttc cct acg cag cag 318
Phe Ala Ala Glu Asp Asp His His Arg His Ser Phe Pro Thr Gln Gln
      40              45              50

```

```

ccg cct cca tcg tcg tcg tcc tca tct ctt atc tca ggt ttc agt aga 366
Pro Pro Pro Ser Ser Ser Ser Ser Ser Leu Ile Ser Gly Phe Ser Arg
      55              60              65

```

```

gag atg gag atg tct gct att gtc tct gct ttg act cac gtt gtt gct 414
Glu Met Glu Met Ser Ala Ile Val Ser Ala Leu Thr His Val Val Ala
      70              75              80              85

```

```

gga aat gtt cct cag cat caa caa gga ggc ggt gaa ggt agc gga gaa 462
Gly Asn Val Pro Gln His Gln Gln Gly Gly Glu Gly Ser Gly Glu
      90              95              100

```

```

ggg act tcg aat tcg tct tct tcc tcg ggg cag aaa agg agg aga gag 510
Gly Thr Ser Asn Ser Ser Ser Ser Ser Gly Gln Lys Arg Arg Arg Glu

```

105	110	115	
gtg gag gaa ggt ggc gcc aaa gcg gtt aag gca gct aat act ttg acg	558		
Val Glu Gly Gly Ala Lys Ala Val Lys Ala Ala Asn Thr Leu Thr			
120	125	130	
ggt gat caa tat ttc tcc ggt ggt agc tct act tct aaa gtg aga gaa	606		
Val Asp Gln Tyr Phe Ser Gly Gly Ser Ser Thr Ser Lys Val Arg Glu			
135	140	145	
gct tcg agt aac atg tca ggt ccg ggc cca aca tac gag tat aca act	654		
Ala Ser Ser Asn Met Ser Gly Pro Gly Pro Thr Tyr Glu Tyr Thr Thr			
150	155	160	165
acg gca act gct agt agc gaa acg tcg tcg ttt agt ggg gac caa cct	702		
Thr Ala Thr Ala Ser Ser Glu Thr Ser Ser Phe Ser Gly Asp Gln Pro			
170	175	180	
cgg cga aga tac aga gga gtt aga caa aga cca tgg gga aag tgg gcg	750		
Arg Arg Arg Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Trp Ala			
185	190	195	
gct gag att cga gat cca ttt aaa gca gct aga gtt tgg ctc ggt acg	798		
Ala Glu Ile Arg Asp Pro Phe Lys Ala Ala Arg Val Trp Leu Gly Thr			
200	205	210	
ttc gac aat gct gaa tca gca gca aga gct tac gac gaa gct gca ctt	846		
Phe Asp Asn Ala Glu Ser Ala Ala Arg Ala Tyr Asp Glu Ala Ala Leu			
215	220	225	
cgg ttt aga ggc aac aaa gcc aaa ctc aac ttc cct gaa aac gtc aaa	894		
Arg Phe Arg Gly Asn Lys Ala Lys Leu Asn Phe Pro Glu Asn Val Lys			
230	235	240	245
ctc gtt aga cct gct tca acc gaa gca caa cct gtg cac caa acc gct	942		
Leu Val Arg Pro Ala Ser Thr Glu Ala Gln Pro Val His Gln Thr Ala			
250	255	260	
gct caa aga ccg acc cag tca agg aac tcg ggt tca acg act acc ctt	990		
Ala Gln Arg Pro Thr Gln Ser Arg Asn Ser Gly Ser Thr Thr Thr Leu			
265	270	275	
ttg ccc ata aga cct gct tcg aat caa agc gtt cat tcg cag ccg ttg	1038		
Leu Pro Ile Arg Pro Ala Ser Asn Gln Ser Val His Ser Gln Pro Leu			
280	285	290	
atg caa tca tac aac ttg agt tac tct gaa atg gct cgt caa caa caa	1086		
Met Gln Ser Tyr Asn Leu Ser Tyr Ser Glu Met Ala Arg Gln Gln Gln			
295	300	305	
cag ttt cag caa cat cat caa caa tct ttg gat tta tac gat caa atg	1134		
Gln Phe Gln Gln His His Gln Gln Ser Leu Asp Leu Tyr Asp Gln Met			
310	315	320	325
tcg ttt ccg ttg cgt ttc ggt cac act gga ggt tca atg atg caa tct	1182		
Ser Phe Pro Leu Arg Phe Gly His Thr Gly Gly Ser Met Met Gln Ser			
330	335	340	

acg tcg tca tca tca tct cat tct cgt cct ctg ttt tcc ccg gct gct 1230  
 Thr Ser Ser Ser Ser Ser His Ser Arg Pro Leu Phe Ser Pro Ala Ala  
                   345                  350                  355

gtt .cag ccg cca cca gaa tca gct agc gaa acc ggt tat ctg cag gat 1278  
 Val Gln Pro Pro Pro Glu Ser Ala Ser Glu Thr Gly Tyr Leu Gln Asp  
                   360                  365                  370

ata caa tgg cca tca gac aag act agt aat aac tac aat aat agt cca 1326  
 Ile Gln Trp Pro Ser Asp Lys Thr Ser Asn Asn Tyr Asn Asn Ser Pro  
                   375                  380                  385

tcc tcc tga tgacttgctt cattttattt gtttcactat agagtaatat 1375  
 Ser Ser \*  
 390

aaaacaggaa aatgattata tggtatagag ttatttttcc aaatattata gggtttaggt 1435  
 tggttgattt gttctgcttt catcctctca tgcttttttt cttaatttat tatatttttg 1495  
 cattataatt tcgtttcatt gtaacaaaca ttaaaaagac cacatggaga aaggaaaaaa 1555  
 aagagag 1562

<210> 48

<211> 391

<212> PRT

<213> Arabidopsis thaliana

<400> 48

Met Cys Val Leu Lys Val Ala Asn Gln Glu Asp Asn Val Gly Lys Lys  
 1                  5                  10                  15  
 Ala Glu Ser Ile Arg Asp Asp Asp His Arg Thr Leu Ser Glu Ile Asp  
                   20                  25                  30  
 Gln Trp Leu Tyr Leu Phe Ala Ala Glu Asp Asp His His Arg His Ser  
                   35                  40                  45  
 Phe Pro Thr Gln Gln Pro Pro Pro Ser Ser Ser Ser Ser Ser Leu Ile  
                   50                  55                  60  
 Ser Gly Phe Ser Arg Glu Met Glu Met Ser Ala Ile Val Ser Ala Leu  
 65                  70                  75                  80  
 Thr His Val Val Ala Gly Asn Val Pro Gln His Gln Gln Gly Gly Gly  
                   85                  90                  95  
 Glu Gly Ser Gly Glu Gly Thr Ser Asn Ser Ser Ser Ser Ser Gly Gln  
                   100                  105                  110  
 Lys Arg Arg Arg Glu Val Glu Glu Gly Gly Ala Lys Ala Val Lys Ala  
                   115                  120                  125  
 Ala Asn Thr Leu Thr Val Asp Gln Tyr Phe Ser Gly Gly Ser Ser Thr  
                   130                  135                  140  
 Ser Lys Val Arg Glu Ala Ser Ser Asn Met Ser Gly Pro Gly Pro Thr  
 145                  150                  155                  160  
 Tyr Glu Tyr Thr Thr Thr Ala Thr Ala Ser Ser Glu Thr Ser Ser Phe  
                   165                  170                  175  
 Ser Gly Asp Gln Pro Arg Arg Arg Tyr Arg Gly Val Arg Gln Arg Pro  
                   180                  185                  190  
 Trp Gly Lys Trp Ala Ala Glu Ile Arg Asp Pro Phe Lys Ala Ala Arg  
                   195                  200                  205  
 Val Trp Leu Gly Thr Phe Asp Asn Ala Glu Ser Ala Ala Arg Ala Tyr  
                   210                  215                  220  
 Asp Glu Ala Ala Leu Arg Phe Arg Gly Asn Lys Ala Lys Leu Asn Phe



225					230					235					240
Pro	Glu	Asn	Val	Lys	Leu	Val	Arg	Pro	Ala	Ser	Thr	Glu	Ala	Gln	Pro
				245					250					255	
Val	His	Gln	Thr	Ala	Ala	Gln	Arg	Pro	Thr	Gln	Ser	Arg	Asn	Ser	Gly
				260					265					270	
Ser	Thr	Thr	Thr	Leu	Leu	Pro	Ile	Arg	Pro	Ala	Ser	Asn	Gln	Ser	Val
				275					280					285	
His	Ser	Gln	Pro	Leu	Met	Gln	Ser	Tyr	Asn	Leu	Ser	Tyr	Ser	Glu	Met
				290					295					300	
Ala	Arg	Gln	Gln	Gln	Gln	Phe	Gln	Gln	His	His	Gln	Gln	Ser	Leu	Asp
305					310					315					320
Leu	Tyr	Asp	Gln	Met	Ser	Phe	Pro	Leu	Arg	Phe	Gly	His	Thr	Gly	Gly
				325					330					335	
Ser	Met	Met	Gln	Ser	Thr	Ser	Ser	Ser	Ser	Ser	His	Ser	Arg	Pro	Leu
				340					345					350	
Phe	Ser	Pro	Ala	Ala	Val	Gln	Pro	Pro	Pro	Glu	Ser	Ala	Ser	Glu	Thr
				355					360					365	
Gly	Tyr	Leu	Gln	Asp	Ile	Gln	Trp	Pro	Ser	Asp	Lys	Thr	Ser	Asn	Asn
				370					375					380	
Tyr	Asn	Asn	Ser	Pro	Ser	Ser									
385					390										

```
<210> 49
<211> 1605
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>
<221> CDS
<222> (90) ... (1547)
```

```

<400> 49
tctcacctat tgtaaaaatc accagtttcg tatataaaac cctaattttc tcaaaattcc 60
caaatattga cttggaatca aaaaatccga atg gat gtg agc aaa gta acc aca 113
Met Asp Val Ser Lys Val Thr Thr
1 5

```

agc gac ggc gga gga gat tca atg gag act aag cca tct cct caa cct 161  
 Ser Asp Gly Gly Gly Asp Ser Met Glu Thr Lys Pro Ser Pro Gln Pro  
           10                          15                          20

cag cct gcg gcg att cta agt tca aac gcg cct cct ccg ttt ctg agc 209  
Gln Pro Ala Ala Ile Leu Ser Ser Asn Ala Pro Pro Pro Phe Leu Ser  
25 30 35 40

aag acc tat gat atg gtt gat gat cac aat aca gat tcg att gtc tct 257  
Lys Thr Tyr Asp Met Val Asp Asp His Asn Thr Asp Ser Ile Val Ser  
45 50 55

tgg agt gct aat aac aac agt ttt atc gtt tgg aaa cca ccg gag ttc 305  
Trp Ser Ala Asn Asn Asn Ser Phe Ile Val Trp Lys Pro Pro Glu Phe  
60 65 70

gct cgc gat. ctt ctt cct aag aac ttt aag cat aat aat ttc tcc agc 353  
Ala Arg Asp Leu Leu Pro Lys Asn Phe Lys His Asn Asn Phe Ser Ser  
75 80 85

ttc gtt aga cag ctt aat acc tat ggt ttc agg aag gtt gac cca gat Phe Val Arg Gln Leu Asn Thr Tyr Gly Phe Arg Lys Val Asp Pro Asp 90 95 100	401
aga tgg gaa ttt gcg aat gaa ggt ttt tta aga ggt cag aag cac ttg Arg Trp Glu Phe Ala Asn Glu Gly Phe Leu Arg Gly Gln Lys His Leu 105 110 115 120	449
cta caa tca ata act agg cga aaa cct gcc cat gga cag gga cag gga Leu Gln Ser Ile Thr Arg Arg Lys Pro Ala His Gly Gln Gly Gln Gly 125 130 135	497
cat cag cga tct cag cac tcg aat gga cag aac tca tct gtt agc gca His Gln Arg Ser Gln His Ser Asn Gly Gln Asn Ser Ser Val Ser Ala 140 145 150	545
tgt gtt gaa gtt ggc aaa ttt ggt ctc gaa gaa gaa gtt gaa agg ctt Cys Val Glu Val Gly Lys Phe Gly Leu Glu Glu Glu Val Glu Arg Leu 155 160 165	593
aaa aga gat aag aac gtc ctt atg caa gaa ctc gtc aga tta aga cag Lys Arg Asp Lys Asn Val Leu Met Gln Glu Leu Val Arg Leu Arg Gln 170 175 180	641
cag caa cag tcc act gat aac caa ctt caa acg atg gtt cag cgt ctc Gln Gln Gln Ser Thr Asp Asn Gln Leu Gln Thr Met Val Gln Arg Leu 185 190 195 200	689
cag ggc atg gag aat cgg caa caa caa tta atg tca ttc ctt gca aag Gln Gly Met Glu Asn Arg Gln Gln Gln Leu Met Ser Phe Leu Ala Lys 205 210 215	737
gca gta caa agc cct cat ttt cta tct caa ttc tta cag cag cag aat Ala Val Gln Ser Pro His Phe Leu Ser Gln Phe Leu Gln Gln Gln Asn 220 225 230	785
cag caa aac gag agt aat agg cgc atc agt gat acc agt aag aag cgg Gln Gln Asn Glu Ser Asn Arg Arg Ile Ser Asp Thr Ser Lys Lys Arg 235 240 245	833
aga ttc aag cga gac ggc att gtc cgt aat aat gat tct gct act cct Arg Phe Lys Arg Asp Gly Ile Val Arg Asn Asn Asp Ser Ala Thr Pro 250 255 260	881
gat gga cag ata gtg aag tat caa cct cca atg cac gag caa gcc aaa Asp Gly Gln Ile Val Lys Tyr Gln Pro Pro Met His Glu Gln Ala Lys 265 270 275 280	929
gca atg ttt aaa cag ctt atg aag atg gaa cct tac aaa acc ggc gat Ala Met Phe Lys Gln Leu Met Lys Met Glu Pro Tyr Lys Thr Gly Asp 285 290 295	977
gat ggt ttc ctt cta ggt aat ggt acg tct act acc gag gga aca gag Asp Gly Phe Leu Leu Gly Asn Gly Thr Ser Thr Thr Glu Gly Thr Glu 300 305 310	1025

atg gag act tca tca aac caa gta tcg ggt ata act ctt aag gaa atg 1073  
 Met Glu Thr Ser Ser Asn Gln Val Ser Gly Ile Thr Leu Lys Glu Met  
 315 320 325

cct aca gct tct gag ata cag tca tca tca cca att gaa aca act cct 1121  
 Pro Thr Ala Ser Glu Ile Gln Ser Ser Ser Pro Ile Glu Thr Thr Pro  
 330 335 340

gaa aat gtt tcg gca gca tca gaa gca acc gag aac tgt att cct tca 1169  
 Glu Asn Val Ser Ala Ala Ser Glu Ala Thr Glu Asn Cys Ile Pro Ser  
 345 350 355 360

cct gat gat cta act ctt ccc gac ttc act cat atg cta ccg gaa aat 1217  
 Pro Asp Asp Leu Thr Leu Pro Asp Phe Thr His Met Leu Pro Glu Asn  
 365 370 375

aat tca gag aag cct cca gag agt ttc atg gaa cca aac ctg gga ggt 1265  
 Asn Ser Glu Lys Pro Pro Glu Ser Phe Met Glu Pro Asn Leu Gly Gly  
 380 385 390

tct agt cca tta cta gat cca gat ctg ttg atc gat gat tct ttg tcc 1313  
 Ser Ser Pro Leu Leu Asp Pro Asp Leu Leu Ile Asp Asp Ser Leu Ser  
 395 400 405

ttc gac att gac gac ttt cca atg gat tct gat ata gac cct gtt gat 1361  
 Phe Asp Ile Asp Asp Phe Pro Met Asp Ser Asp Ile Asp Pro Val Asp  
 410 415 420

tac ggt tta ctc gaa cgc tta ctc atg tca agc ccg gtt cca gat aat 1409  
 Tyr Gly Leu Leu Glu Arg Leu Leu Met Ser Ser Pro Val Pro Asp Asn  
 425 430 435 440

atg gat tca aca cca gtg gac aat gaa aca gag cag gaa caa aat gga 1457  
 Met Asp Ser Thr Pro Val Asp Asn Glu Thr Glu Gln Glu Gln Asn Gly  
 445 450 455

tgg gac aaa act aag cat atg gat aat ctg act caa cag atg ggt ctc 1505  
 Trp Asp Lys Thr Lys His Met Asp Asn Leu Thr Gln Gln Met Gly Leu  
 460 465 470

ctc tct cct gaa acc tta gat ctc tca agg caa aat cct tga 1547  
 Leu Ser Pro Glu Thr Leu Asp Leu Ser Arg Gln Asn Pro \*  
 475 480 485

ttttgggagt ttttaaagtc ttttgaggta acacagtccc tgagagcagc atattcat 1605

<210> 50  
 <211> 485  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (32)...(130)  
 <223> Conserved domain

<400> 50

Met Asp Val Ser Lys Val Thr Thr Ser Asp Gly Gly Gly Asp Ser Met  
 1 5 10 15  
 Glu Thr Lys Pro Ser Pro Gln Pro Gln Pro Ala Ala Ile Leu Ser Ser  
 20 25 30  
 Asn Ala Pro Pro Phe Leu Ser Lys Thr Tyr Asp Met Val Asp Asp  
 35 40 45  
 His Asn Thr Asp Ser Ile Val Ser Trp Ser Ala Asn Asn Ser Phe  
 50 55 60  
 Ile Val Trp Lys Pro Pro Glu Phe Ala Arg Asp Leu Leu Pro Lys Asn  
 65 70 75 80  
 Phe Lys His Asn Asn Phe Ser Ser Phe Val Arg Gln Leu Asn Thr Tyr  
 85 90 95  
 Gly Phe Arg Lys Val Asp Pro Asp Arg Trp Glu Phe Ala Asn Glu Gly  
 100 105 110  
 Phe Leu Arg Gly Gln Lys His Leu Leu Gln Ser Ile Thr Arg Arg Lys  
 115 120 125  
 Pro Ala His Gly Gln Gly Gln Gly His Gln Arg Ser Gln His Ser Asn  
 130 135 140  
 Gly Gln Asn Ser Ser Val Ser Ala Cys Val Glu Val Gly Lys Phe Gly  
 145 150 155 160  
 Leu Glu Glu Glu Val Glu Arg Leu Lys Arg Asp Lys Asn Val Leu Met  
 165 170 175  
 Gln Glu Leu Val Arg Leu Arg Gln Gln Gln Gln Ser Thr Asp Asn Gln  
 180 185 190  
 Leu Gln Thr Met Val Gln Arg Leu Gln Gly Met Glu Asn Arg Gln Gln  
 195 200 205  
 Gln Leu Met Ser Phe Leu Ala Lys Ala Val Gln Ser Pro His Phe Leu  
 210 215 220  
 Ser Gln Phe Leu Gln Gln Gln Asn Gln Gln Asn Glu Ser Asn Arg Arg  
 225 230 235 240  
 Ile Ser Asp Thr Ser Lys Lys Arg Arg Phe Lys Arg Asp Gly Ile Val  
 245 250 255  
 Arg Asn Asn Asp Ser Ala Thr Pro Asp Gly Gln Ile Val Lys Tyr Gln  
 260 265 270  
 Pro Pro Met His Glu Gln Ala Lys Ala Met Phe Lys Gln Leu Met Lys  
 275 280 285  
 Met Glu Pro Tyr Lys Thr Gly Asp Asp Gly Phe Leu Leu Gly Asn Gly  
 290 295 300  
 Thr Ser Thr Thr Glu Gly Thr Glu Met Glu Thr Ser Ser Asn Gln Val  
 305 310 315 320  
 Ser Gly Ile Thr Leu Lys Glu Met Pro Thr Ala Ser Glu Ile Gln Ser  
 325 330 335  
 Ser Ser Pro Ile Glu Thr Thr Pro Glu Asn Val Ser Ala Ala Ser Glu  
 340 345 350  
 Ala Thr Glu Asn Cys Ile Pro Ser Pro Asp Asp Leu Thr Leu Pro Asp  
 355 360 365  
 Phe Thr His Met Leu Pro Glu Asn Asn Ser Glu Lys Pro Pro Glu Ser  
 370 375 380  
 Phe Met Glu Pro Asn Leu Gly Gly Ser Ser Pro Leu Leu Asp Pro Asp  
 385 390 395 400  
 Leu Leu Ile Asp Asp Ser Leu Ser Phe Asp Ile Asp Asp Phe Pro Met  
 405 410 415  
 Asp Ser Asp Ile Asp Pro Val Asp Tyr Gly Leu Leu Glu Arg Leu Leu  
 420 425 430  
 Met Ser Ser Pro Val Pro Asp Asn Met Asp Ser Thr Pro Val Asp Asn  
 435 440 445  
 Glu Thr Glu Gln Glu Gln Asn Gly Trp Asp Lys Thr Lys His Met Asp

**<220>**

&lt;221&gt; DOMAIN

&lt;222&gt; (39)...(76)

&lt;223&gt; Conserved domain

&lt;400&gt; 52

```

Met Phe Arg Ser Asp Lys Ala Glu Lys Met Asp Lys Arg Arg Arg Arg
 1           5           10           15
Gln Ser Lys Ala Lys Ala Ser Cys Ser Glu Glu Val Ser Ser Ile Glu
          20           25           30
Trp Glu Ala Val Lys Met Ser Glu Glu Glu Glu Asp Leu Ile Ser Arg
          35           40           45
Met Tyr Lys Leu Val Gly Asp Arg Trp Glu Leu Ile Ala Gly Arg Ile
          50           55           60
Pro Gly Arg Thr Pro Glu Glu Ile Glu Arg Tyr Trp Leu Met Lys His
65           70           75           80
Gly Val Val Phe Ala Asn Arg Arg Arg Asp Phe Phe Arg Lys
          85           90

```

&lt;210&gt; 53

&lt;211&gt; 407

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (10)...(348)

&lt;400&gt; 53

```

ccagtagtt atg gat aat acc aac cgt ctt cgt ctt cgt cgc ggt ccc agt 51
      Met Asp Asn Thr Asn Arg Leu Arg Leu Arg Arg Gly Pro Ser
        1           5           10

ctt agg caa act aag ttc act cga tcc cga tat gac tct gaa gaa gtg   99
Leu Arg Gln Thr Lys Phe Thr Arg Ser Arg Tyr Asp Ser Glu Glu Val
 15           20           25           30

agt agc atc gaa tgg gag ttt atc agt atg acc gaa caa gaa gaa gat   147
Ser Ser Ile Glu Trp Glu Phe Ile Ser Met Thr Glu Gln Glu Glu Asp
          35           40           45

ctc atc tct cga atg tac aga ctt gtc ggt aat agg tgg gat tta ata   195
Leu Ile Ser Arg Met Tyr Arg Leu Val Gly Asn Arg Trp Asp Leu Ile
          50           55           60

gca gga aga gtc gta gga aga aag gca aat gag att gag aga tac tgg   243
Ala Gly Arg Val Val Gly Arg Lys Ala Asn Glu Ile Glu Arg Tyr Trp
          65           70           75

att atg aga aac tct gac tat ttt tct cac aaa cga cga cgt ctt aat   291
Ile Met Arg Asn Ser Asp Tyr Phe Ser His Lys Arg Arg Arg Leu Asn
          80           85           90

aat tct ccc ttt ttt tct act tct cct ctt aat ctc caa gaa aat cta   339
Asn Ser Pro Phe Phe Ser Thr Ser Pro Leu Asn Leu Gln Glu Asn Leu
          95           100           105           110

```

aaa ttg taa agaaatcaaa ataaaagctt tcaatcataa aagtagaaca 388  
Lys Leu \*

aatcttgaat gtcttctca 407

<210> 54  
<211> 112  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<221> DOMAIN  
<222> (28)...(78)  
<223> Conserved domain

<400> 54  
Met Asp Asn Thr Asn Arg Leu Arg Leu Arg Arg Gly Pro Ser Leu Arg  
1 5 10 15  
Gln Thr Lys Phe Thr Arg Ser Arg Tyr Asp Ser Glu Glu Val Ser Ser  
20 25 30  
Ile Glu Trp Glu Phe Ile Ser Met Thr Glu Gln Glu Asp Leu Ile  
35 40 45  
Ser Arg Met Tyr Arg Leu Val Gly Asn Arg Trp Asp Leu Ile Ala Gly  
50 55 60  
Arg Val Val Gly Arg Lys Ala Asn Glu Ile Glu Arg Tyr Trp Ile Met  
65 70 75 80  
Arg Asn Ser Asp Tyr Phe Ser His Lys Arg Arg Arg Leu Asn Asn Ser  
85 90 95  
Pro Phe Phe Ser Thr Ser Pro Leu Asn Leu Gln Glu Asn Leu Lys Leu  
100 105 110

<210> 55  
<211> 2405  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (381)...(2213)

<400> 55  
aagccacaca atctcttttc ttctctctct ctctgttata tctcttctgt ttaattcttt 60  
tattcttctt cgtctatctt ctctataat ctcttctctc tccctcttca cctaaagaat 120  
aagaagaaaa ataattcaca tctttatgca aactactttc ttgtagggtt ttaggagcta 180  
tctctattgt cttggttctg atacaaagtt ttgtaatttt catgggtatga gaagatttgc 240  
ctttctattt tgtttattgg ttctttttaa ctttttcttg gagatgggtt cttgtagatc 300  
ttaatgaaac ttctgttttt gtcccaaaaa gagtttctt ttttcttctc ttctttttgg 360  
gttttcaatt cttgagagac atg gca aga gat cag ttc tat ggt cac aat aac 413  
Met Ala Arg Asp Gln Phe Tyr Gly His Asn Asn  
1 5 10

cat cat cat caa gag caa caa cat caa atg att aat cag atc caa ggg 461  
His His His Gln Gln Gln His Gln Met Ile Asn Gln Ile Gln Gly  
15 20 25

ttt gat gag aca aac caa aac cca acc gat cat cat cat tac aat cat	509
Phe Asp Glu Thr Asn Gln Asn Pro Thr Asp His His Tyr Asn His	
30 35 40	
cag atc ttt ggc tca aac tcc aac atg ggt atg atg ata gac ttc tct	557
Gln Ile Phe Gly Ser Asn Ser Asn Met Gly Met Met Ile Asp Phe Ser	
45 50 55	
aag caa caa cag att agg atg aca agt ggt tgc gat cat cat cat cat	605
Lys Gln Gln Gln Ile Arg Met Thr Ser Gly Ser Asp His His His His	
60 65 70 75	
cat cat cag aca agt ggt ggt act gat cag aat cag ctt ctg gaa gat	653
His His Gln Thr Ser Gly Gly Thr Asp Gln Asn Gln Leu Leu Glu Asp	
80 85 90	
tct tca tct gcc atg aga cta tgc aat gtt aat aat gat ttc cca agt	701
Ser Ser Ser Ala Met Arg Leu Cys Asn Val Asn Asn Asp Phe Pro Ser	
95 100 105	
gaa gta aat gat gag aga cca cca caa aga cca agc caa ggt ctt tcc	749
Glu Val Asn Asp Glu Arg Pro Pro Gln Arg Pro Ser Gln Gly Leu Ser	
110 115 120	
ctt tct ctc tcc tct tca aat cct aca agc atc agt ctc caa tct ttc	797
Leu Ser Leu Ser Ser Ser Asn Pro Thr Ser Ile Ser Leu Gln Ser Phe	
125 130 135	
gaa ctc aga ccc caa caa caa caa caa ggg tat tcc ggt aat aaa tca	845
Glu Leu Arg Pro Gln Gln Gln Gln Gly Tyr Ser Gly Asn Lys Ser	
140 145 150 155	
aca caa cat cag aat ctc caa cac acg cag atg atg atg atg atg atg	893
Thr Gln His Gln Asn Leu Gln His Thr Gln Met Met Met Met Met Met	
160 165 170	
aat agt cac cac caa aac aac aac aat aac aat cat cag cat cat aat	941
Asn Ser His His Gln Asn Asn Asn Asn Asn Asn His Gln His His Asn	
175 180 185	
cat cat cag ttt cag att ggg agt tcc aag tat ttg agt cca gct caa	989
His His Gln Phe Gln Ile Gly Ser Ser Lys Tyr Leu Ser Pro Ala Gln	
190 195 200	
gag cta ctg agt gag ttt tgc agt ctt gga gta aag gaa agc gat gaa	1037
Glu Leu Leu Ser Glu Phe Cys Ser Leu Gly Val Lys Glu Ser Asp Glu	
205 210 215	
gaa gtg atg atg atg aag cat aag aag aag caa aag ggt aaa caa caa	1085
Glu Val Met Met Met Lys His Lys Lys Lys Gln Lys Gly Lys Gln Gln	
220 225 230 235	
gaa gag tgg gac aca agt cac cac agc aac aat gat caa cat gac caa	1133
Glu Glu Trp Asp Thr Ser His His Ser Asn Asn Asp Gln His Asp Gln	
240 245 250	
tct gcg act act tct tca aag aaa cat gtt cca cca ctt cac tct ctt	1181



Ser	Ala	Thr	Thr	Ser	Ser	Lys	Lys	His	Val	Pro	Pro	Leu	His	Ser	Leu		
			255					260					265				
gag	ttc	atg	gaa	ctt	cag	aaa	aga	aaa	gcc	aag	ttg	ctc	tcc	atg	ctc	1229	
Glu	Phe	Met	Glu	Leu	Gln	Lys	Arg	Lys	Ala	Lys	Leu	Leu	Ser	Met	Leu		
		270					275					280					
gaa	gag	ctt	aaa	aga	aga	tat	gga	cat	tac	cga	gag	caa	atg	aga	gtt	1277	
Glu	Glu	Leu	Lys	Arg	Arg	Tyr	Gly	His	Tyr	Arg	Glu	Gln	Met	Arg	Val		
		285				290					295						
gcg	gcg	gca	gcc	ttt	gaa	gcg	gcg	gtt	gga	cta	gga	ggg	gca	gag	ata	1325	
Ala	Ala	Ala	Ala	Phe	Glu	Ala	Ala	Val	Gly	Leu	Gly	Gly	Ala	Glu	Ile		
300					305				310						315		
tac	act	gcg	tta	gcg	tca	agg	gca	atg	tca	aga	cac	ttt	cgg	tgt	tta	1373	
Tyr	Thr	Ala	Leu	Ala	Ser	Arg	Ala	Met	Ser	Arg	His	Phe	Arg	Cys	Leu		
				320					325					330			
aaa	gac	gga	ctt	gtg	gga	cag	att	caa	gca	aca	agt	caa	gct	ttg	gga	1421	
Lys	Asp	Gly	Leu	Val	Gly	Gln	Ile	Gln	Ala	Thr	Ser	Gln	Ala	Leu	Gly		
			335				340					345					
gag	aga	gaa	gag	gat	aat	cgt	gcg	gtt	tct	att	gca	gca	cgt	gga	gaa	1469	
Glu	Arg	Glu	Glu	Asp	Asn	Arg	Ala	Val	Ser	Ile	Ala	Ala	Arg	Gly	Glu		
		350					355					360					
act	cca	cgg	ttg	aga	ttg	ctc	gat	caa	gct	ttg	cgg	caa	cag	aaa	tcg	1517	
Thr	Pro	Arg	Leu	Arg	Leu	Leu	Asp	Gln	Ala	Leu	Arg	Gln	Gln	Lys	Ser		
		365				370				375							
tat	cgc	caa	atg	act	ctt	gtt	gac	gct	cat	cct	tgg	cgt	cca	caa	cgc	1565	
Tyr	Arg	Gln	Met	Thr	Leu	Val	Asp	Ala	His	Pro	Trp	Arg	Pro	Gln	Arg		
380					385					390					395		
ggc	ttg	cct	gaa	cgc	gca	gtc	aca	acg	ttg	aga	gct	tgg	ctc	ttt	gaa	1613	
Gly	Leu	Pro	Glu	Arg	Ala	Val	Thr	Thr	Leu	Arg	Ala	Trp	Leu	Phe	Glu		
			400						405					410			
cac	ttt	ctt	cac	cca	tat	ccg	agc	gat	gtt	gat	aag	cat	ata	ttg	gcc	1661	
His	Phe	Leu	His	Pro	Tyr	Pro	Ser	Asp	Val	Asp	Lys	His	Ile	Leu	Ala		
			415					420					425				
cga	caa	act	ggt	tta	tca	aga	agt	cag	gta	tca	aat	tgg	ttt	att	aat	1709	
Arg	Gln	Thr	Gly	Leu	Ser	Arg	Ser	Gln	Val	Ser	Asn	Trp	Phe	Ile	Asn		
		430					435					440					
gca	aga	gtt	agg	cta	tgg	aaa	cca	atg	att	gaa	gaa	atg	tac	tgt	gaa	1757	
Ala	Arg	Val	Arg	Leu	Trp	Lys	Pro	Met	Ile	Glu	Glu	Met	Tyr	Cys	Glu		
		445				450					455						
gaa	aca	aga	agt	gaa	caa	atg	gag	att	aca	aac	ccg	atg	atg	atc	gat	1805	
Glu	Thr	Arg	Ser	Glu	Gln	Met	Glu	Ile	Thr	Asn	Pro	Met	Met	Ile	Asp		
460					465					470				475			
act	aaa	ccg	gac	ccg	gac	cag	ttg	atc	cgt	gtc	gaa	ccg	gaa	tct	tta	1853	
Thr	Lys	Pro	Asp	Pro	Asp	Gln	Leu	Ile	Arg	Val	Glu	Pro	Glu	Ser	Leu		

```

          480          485          490
tcc tca ata gtg aca aac cct aca tcc aaa tcc ggt cac aac tca acc 1901
Ser Ser Ile Val Thr Asn Pro Thr Ser Lys Ser Gly His Asn Ser Thr
          495          500          505

cat gga acg atg tcg tta ggg tca acg ttt gac ttt tcc ttg tac ggt 1949
His Gly Thr Met Ser Leu Gly Ser Thr Phe Asp Phe Ser Leu Tyr Gly
          510          515          520

aac caa gct gtg aca tac gct ggt gaa gga ggg cca cgt ggt gac gtt 1997
Asn Gln Ala Val Thr Tyr Ala Gly Glu Gly Gly Pro Arg Gly Asp Val
          525          530          535

tcc ttg acg ctt ggg tta caa cgt aac gat ggt aac ggt ggt gtg agt 2045
Ser Leu Thr Leu Gly Leu Gln Arg Asn Asp Gly Asn Gly Gly Val Ser
          540          545          550          555

tta gcg ttg tct cca gtg acg gct caa ggt ggc caa ctt ttc tac ggt 2093
Leu Ala Leu Ser Pro Val Thr Ala Gln Gly Gly Gln Leu Phe Tyr Gly
          560          565          570

aga gac cac att gaa gaa gga ccg gtt caa tat tca gcg tcg atg tta 2141
Arg Asp His Ile Glu Glu Gly Pro Val Gln Tyr Ser Ala Ser Met Leu
          575          580          585

gat gat gat caa gtt cag aat ttg cct tat agg aat ttg atg gga gct 2189
Asp Asp Asp Gln Val Gln Asn Leu Pro Tyr Arg Asn Leu Met Gly Ala
          590          595          600

caa tta ctt cat gat att gtt tga gattaaaaga ttaggaccaa agttatcgat 2243
Gln Leu Leu His Asp Ile Val *
          605          610

acatatttttc caaaaccgat tcggttatgt aacgggttag ttagataaaa accaaattag 2303
atatttatat ataccgttgt ctgattggat tggaggattg gtggacaagg agatattatt 2363
aatgtatgag ttagtttggt cgtcaaaaaa aaaaaaaaaa aa 2405

```

&lt;210&gt; 56

&lt;211&gt; 610

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (392)...(452)

&lt;223&gt; Conserved domain

&lt;400&gt; 56

```

Met Ala Arg Asp Gln Phe Tyr Gly His Asn Asn His His His Gln Glu
1      5      10      15
Gln Gln His Gln Met Ile Asn Gln Ile Gln Gly Phe Asp Glu Thr Asn
20     25     30
Gln Asn Pro Thr Asp His His His Tyr Asn His Gln Ile Phe Gly Ser
35     40     45
Asn Ser Asn Met Gly Met Met Ile Asp Phe Ser Lys Gln Gln Gln Ile
50     55     60

```

Arg Met Thr Ser Gly Ser Asp His His His His His His Gln Thr Ser  
 65 70 75 80  
 Gly Gly Thr Asp Gln Asn Gln Leu Leu Glu Asp Ser Ser Ser Ala Met  
 85 90 95  
 Arg Leu Cys Asn Val Asn Asn Asp Phe Pro Ser Glu Val Asn Asp Glu  
 100 105 110  
 Arg Pro Pro Gln Arg Pro Ser Gln Gly Leu Ser Leu Ser Leu Ser Ser  
 115 120 125  
 Ser Asn Pro Thr Ser Ile Ser Leu Gln Ser Phe Glu Leu Arg Pro Gln  
 130 135 140  
 Gln Gln Gln Gln Gly Tyr Ser Gly Asn Lys Ser Thr Gln His Gln Asn  
 145 150 155 160  
 Leu Gln His Thr Gln Met Met Met Met Met Met Asn Ser His His Gln  
 165 170 175  
 Asn Asn Asn Asn Asn Asn His Gln His His Asn His His Gln Phe Gln  
 180 185 190  
 Ile Gly Ser Ser Lys Tyr Leu Ser Pro Ala Gln Glu Leu Leu Ser Glu  
 195 200 205  
 Phe Cys Ser Leu Gly Val Lys Glu Ser Asp Glu Glu Val Met Met Met  
 210 215 220  
 Lys His Lys Lys Lys Gln Lys Gly Lys Gln Gln Glu Glu Trp Asp Thr  
 225 230 235 240  
 Ser His His Ser Asn Asn Asp Gln His Asp Gln Ser Ala Thr Thr Ser  
 245 250 255  
 Ser Lys Lys His Val Pro Pro Leu His Ser Leu Glu Phe Met Glu Leu  
 260 265 270  
 Gln Lys Arg Lys Ala Lys Leu Leu Ser Met Leu Glu Glu Leu Lys Arg  
 275 280 285  
 Arg Tyr Gly His Tyr Arg Glu Gln Met Arg Val Ala Ala Ala Ala Phe  
 290 295 300  
 Glu Ala Ala Val Gly Leu Gly Gly Ala Glu Ile Tyr Thr Ala Leu Ala  
 305 310 315 320  
 Ser Arg Ala Met Ser Arg His Phe Arg Cys Leu Lys Asp Gly Leu Val  
 325 330 335  
 Gly Gln Ile Gln Ala Thr Ser Gln Ala Leu Gly Glu Arg Glu Glu Asp  
 340 345 350  
 Asn Arg Ala Val Ser Ile Ala Ala Arg Gly Glu Thr Pro Arg Leu Arg  
 355 360 365  
 Leu Leu Asp Gln Ala Leu Arg Gln Gln Lys Ser Tyr Arg Gln Met Thr  
 370 375 380  
 Leu Val Asp Ala His Pro Trp Arg Pro Gln Arg Gly Leu Pro Glu Arg  
 385 390 395 400  
 Ala Val Thr Thr Leu Arg Ala Trp Leu Phe Glu His Phe Leu His Pro  
 405 410 415  
 Tyr Pro Ser Asp Val Asp Lys His Ile Leu Ala Arg Gln Thr Gly Leu  
 420 425 430  
 Ser Arg Ser Gln Val Ser Asn Trp Phe Ile Asn Ala Arg Val Arg Leu  
 435 440 445  
 Trp Lys Pro Met Ile Glu Glu Met Tyr Cys Glu Glu Thr Arg Ser Glu  
 450 455 460  
 Gln Met Glu Ile Thr Asn Pro Met Met Ile Asp Thr Lys Pro Asp Pro  
 465 470 475 480  
 Asp Gln Leu Ile Arg Val Glu Pro Glu Ser Leu Ser Ser Ile Val Thr  
 485 490 495  
 Asn Pro Thr Ser Lys Ser Gly His Asn Ser Thr His Gly Thr Met Ser  
 500 505 510  
 Leu Gly Ser Thr Phe Asp Phe Ser Leu Tyr Gly Asn Gln Ala Val Thr

515 520 525  
 Tyr Ala Gly Glu Gly Gly Pro Arg Gly Asp Val Ser Leu Thr Leu Gly  
 530 535 540  
 Leu Gln Arg Asn Asp Gly Asn Gly Gly Val Ser Leu Ala Leu Ser Pro  
 545 550 555 560  
 Val Thr Ala Gln Gly Gly Gln Leu Phe Tyr Gly Arg Asp His Ile Glu  
 565 570 575  
 Glu Gly Pro Val Gln Tyr Ser Ala Ser Met Leu Asp Asp Asp Gln Val  
 580 585 590  
 Gln Asn Leu Pro Tyr Arg Asn Leu Met Gly Ala Gln Leu Leu His Asp  
 595 600 605  
 Ile Val  
 610

<210> 57  
 <211> 890  
 <212> DNA  
 <213> *Arabidopsis thaliana*

<220>  
 <221> CDS  
 <222> (55)...(738)

<400> 57

gcaaccttca aactaaaact cgagagacaa gaaatcctca gaatctttaa ctta atg 57  
 Met  
 1  
 gcg ctc gag gct ctt aca tca cca aga tta gct tct ccg att cct cct 105  
 Ala Leu Glu Ala Leu Thr Ser Pro Arg Leu Ala Ser Pro Ile Pro Pro  
 5 10 15  
 ttg ttc gaa gat tct tca gtc ttc cat gga gtc gag cac tgg aca aag 153  
 Leu Phe Glu Asp Ser Ser Val Phe His Gly Val Glu His Trp Thr Lys  
 20 25 30  
 ggt aag cga tct aag aga tca aga tcc gat ttc cac cac caa aac ctc 201  
 Gly Lys Arg Ser Lys Arg Ser Arg Ser Asp Phe His His Gln Asn Leu  
 35 40 45  
 act gag gaa gag tat cta gct ttt tgc ctc atg ctt ctc gct cgc gac 249  
 Thr Glu Glu Glu Tyr Leu Ala Phe Cys Leu Met Leu Leu Ala Arg Asp  
 50 55 60 65  
 aac cgt cag cct cct cct cct ccg gcg gtg gag aag ttg agc tac aag 297  
 Asn Arg Gln Pro Pro Pro Pro Pro Ala Val Glu Lys Leu Ser Tyr Lys  
 70 75 80  
 tgt agc gtc tgc gac aag acg ttc tct tct tac caa gct ctc ggt ggt 345  
 Cys Ser Val Cys Asp Lys Thr Phe Ser Ser Tyr Gln Ala Leu Gly Gly  
 85 90 95  
 cac aag gca agc cac cgt aag aac tta tca cag act ctc tcc ggc gga 393  
 His Lys Ala Ser His Arg Lys Asn Leu Ser Gln Thr Leu Ser Gly Gly  
 100 105 110

gga gat gat cat tca acc tcg tcg gcg aca acc aca tcc gcc gtg act 441  
 Gly Asp Asp His Ser Thr Ser Ser Ala Thr Thr Thr Ser Ala Val Thr  
 115 120 125

act gga agt ggg aaa tca cac gtt tgc acc atc tgt aac aag tct ttt 489  
 Thr Gly Ser Gly Lys Ser His Val Cys Thr Ile Cys Asn Lys Ser Phe  
 130 135 140 145

cct tcc ggt caa gct ctc ggc gga cac aag cgg tgc cac tac gaa gga 537  
 Pro Ser Gly Gln Ala Leu Gly Gly His Lys Arg Cys His Tyr Glu Gly  
 150 155 160

aac aac aac atc aac act agt agc gtg tcc aac tcc gaa ggt gcg ggg 585  
 Asn Asn Asn Ile Asn Thr Ser Ser Val Ser Asn Ser Glu Gly Ala Gly  
 165 170 175

tcc act agc cac gtt agc agt agc cac cgt ggg ttt gac ctc aac atc 633  
 Ser Thr Ser His Val Ser Ser Ser His Arg Gly Phe Asp Leu Asn Ile  
 180 185 190

cct ccg atc cct gaa ttc tcg atg gtc aac gga gac gac gaa gtc atg 681  
 Pro Pro Ile Pro Glu Phe Ser Met Val Asn Gly Asp Asp Glu Val Met  
 195 200 205

agc cct atg ccg gcg aag aag cct cgg ttt gac ttt ccg gtc aaa ctt 729  
 Ser Pro Met Pro Ala Lys Lys Pro Arg Phe Asp Phe Pro Val Lys Leu  
 210 215 220 225

caa ctt taa ggaaatttac ttagacgata agatttcggtt tgtatactgt 778  
 Gln Leu \*

tgagagttgt gtaggaattt gttgactgta cataccaaat tggactttga ctgattccaa 838  
 ttctttctgt tctttcattt taaaaattat taaaccgatt ctttaccaca aa 890

<210> 58  
 <211> 227  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (82)...(102)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (136)...(154)  
 <223> Conserved domain

<400> 58  
 Met Ala Leu Glu Ala Leu Thr Ser Pro Arg Leu Ala Ser Pro Ile Pro  
 1 5 10 15  
 Pro Leu Phe Glu Asp Ser Ser Val Phe His Gly Val Glu His Trp Thr  
 20 25 30  
 Lys Gly Lys Arg Ser Lys Arg Ser Arg Ser Asp Phe His His Gln Asn  
 35 40 45  
 Leu Thr Glu Glu Glu Tyr Leu Ala Phe Cys Leu Met Leu Leu Ala Arg

50                      55                      60  
 Asp Asn Arg Gln Pro Pro Pro Pro Ala Val Glu Lys Leu Ser Tyr  
 65                      70                      75                      80  
 Lys Cys Ser Val Cys Asp Lys Thr Phe Ser Ser Tyr Gln Ala Leu Gly  
                     85                      90                      95  
 Gly His Lys Ala Ser His Arg Lys Asn Leu Ser Gln Thr Leu Ser Gly  
                     100                      105                      110  
 Gly Gly Asp Asp His Ser Thr Ser Ser Ala Thr Thr Thr Ser Ala Val  
                     115                      120                      125  
 Thr Thr Gly Ser Gly Lys Ser His Val Cys Thr Ile Cys Asn Lys Ser  
                     130                      135                      140  
 Phe Pro Ser Gly Gln Ala Leu Gly Gly His Lys Arg Cys His Tyr Glu  
 145                      150                      155                      160  
 Gly Asn Asn Asn Ile Asn Thr Ser Ser Val Ser Asn Ser Glu Gly Ala  
                     165                      170                      175  
 Gly Ser Thr Ser His Val Ser Ser Ser His Arg Gly Phe Asp Leu Asn  
                     180                      185                      190  
 Ile Pro Pro Ile Pro Glu Phe Ser Met Val Asn Gly Asp Asp Glu Val  
                     195                      200                      205  
 Met Ser Pro Met Pro Ala Lys Lys Pro Arg Phe Asp Phe Pro Val Lys  
                     210                      215                      220  
 Leu Gln Leu  
 225

&lt;210&gt; 59

&lt;211&gt; 1413

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (86)...(1168)

&lt;400&gt; 59

aatttgtttt tttttctttt gtgggttcaa ttccaattgt tttccctgag actcaagtta 60  
 ctgtgtcatt actctgcatt gagca atg ggt agc aac gaa gaa gga aac ccc 112  
                     Met Gly Ser Asn Glu Glu Gly Asn Pro  
                     1                      5  
  
 act aac aac tct gat aag cca tcg caa gct gct gct cct gag cag agt 160  
 Thr Asn Asn Ser Asp Lys Pro Ser Gln Ala Ala Ala Pro Glu Gln Ser  
 10                      15                      20                      25  
  
 aat gtt cat gtg tat cat cat gac tgg gct gct atg cag gca tat tat 208  
 Asn Val His Val Tyr His His Asp Trp Ala Ala Met Gln Ala Tyr Tyr  
                     30                      35                      40  
  
 ggg cct aga gtt ggt ata cct caa tat tac aac tca aat ttg gcg cct 256  
 Gly Pro Arg Val Gly Ile Pro Gln Tyr Tyr Asn Ser Asn Leu Ala Pro  
                     45                      50                      55  
  
 ggt cat gct cca ccg cct tat atg tgg gcg tct cca tcg cca atg atg 304  
 Gly His Ala Pro Pro Pro Tyr Met Trp Ala Ser Pro Ser Pro Met Met  
                     60                      65                      70  
  
 gct cct tat gga gca cca tat cca cca ttt tgc cct cct ggt gga gtt 352

Ala Pro Tyr Gly Ala Pro Tyr Pro Pro Phe Cys Pro Pro Gly Gly Val	
75 80 85	
tat gct cat cct ggt gtt caa atg ggc tca caa cca caa ggt cct gtt	400
Tyr Ala His Pro Gly Val Gln Met Gly Ser Gln Pro Gln Gly Pro Val	
90 95 100 105	
tct caa tca gca tct gga gtt aca acc cct ttg acc att gat gca cca	448
Ser Gln Ser Ala Ser Gly Val Thr Thr Pro Leu Thr Ile Asp Ala Pro	
110 115 120	
gct aat tca gct gga aac tca gat cat ggg ttc atg aaa aag ctg aaa	496
Ala Asn Ser Ala Gly Asn Ser Asp His Gly Phe Met Lys Lys Leu Lys	
125 130 135	
gag ttc gat gga ctt gca atg tca ata agc aat aac aaa gtt ggg agt	544
Glu Phe Asp Gly Leu Ala Met Ser Ile Ser Asn Asn Lys Val Gly Ser	
140 145 150	
gct gaa cat agc agc agt gaa cat agg agt tct cag agc tcc gag aat	592
Ala Glu His Ser Ser Ser Glu His Arg Ser Ser Gln Ser Ser Glu Asn	
155 160 165	
gat ggc tct agc aat ggt agt gat ggt aat aca act ggg gga gaa caa	640
Asp Gly Ser Ser Asn Gly Ser Asp Gly Asn Thr Thr Gly Gly Glu Gln	
170 175 180 185	
tct agg agg aaa aga agg caa caa aga tca cca agc act ggt gaa aga	688
Ser Arg Arg Lys Arg Arg Gln Gln Arg Ser Pro Ser Thr Gly Glu Arg	
190 195 200	
ccc tca tct caa aac agt ctg cct ctt aga ggt gaa aat gag aaa ccc	736
Pro Ser Ser Gln Asn Ser Leu Pro Leu Arg Gly Glu Asn Glu Lys Pro	
205 210 215	
gat gtg act atg ggg act cct gtt atg ccc aca gca atg agt ttc caa	784
Asp Val Thr Met Gly Thr Pro Val Met Pro Thr Ala Met Ser Phe Gln	
220 225 230	
aac tct gct ggc atg aac ggt gtg cca cag cca tgg aat gaa aaa gag	832
Asn Ser Ala Gly Met Asn Gly Val Pro Gln Pro Trp Asn Glu Lys Glu	
235 240 245	
gtt aaa cga gag aag aga aaa cag tca aac cga gaa tct gct agg agg	880
Val Lys Arg Glu Lys Arg Lys Gln Ser Asn Arg Glu Ser Ala Arg Arg	
250 255 260 265	
tca aga ctg agg aag cag gct gaa aca gaa caa cta tct gtc aaa gtt	928
Ser Arg Leu Arg Lys Gln Ala Glu Thr Glu Gln Leu Ser Val Lys Val	
270 275 280	
gac gca tta gta gct gag aac atg tct ctg agg tct aaa cta ggc cag	976
Asp Ala Leu Val Ala Glu Asn Met Ser Leu Arg Ser Lys Leu Gly Gln	
285 290 295	
cta aac aat gag tct gag aaa cta cgg ctg gag aac gaa gct ata ttg	1024
Leu Asn Asn Glu Ser Glu Lys Leu Arg Leu Glu Asn Glu Ala Ile Leu	

```

      300              305              310
gat caa ctg aaa gcg caa gca aca ggg aaa aca gag aac ctg atc tct 1072
Asp Gln Leu Lys Ala Gln Ala Thr Gly Lys Thr Glu Asn Leu Ile Ser
      315              320              325

cga gtt gat aag aac aac tct gta tca ggt agc aaa act gtg cag cat 1120
Arg Val Asp Lys Asn Asn Ser Val Ser Gly Ser Lys Thr Val Gln His
      330              335              340              345

caa ctg tta aat gca agt ccg ata acc gat cct gtc gcg gct agc tga 1168
Gln Leu Leu Asn Ala Ser Pro Ile Thr Asp Pro Val Ala Ala Ser *
      350              355              360

ccgtggccgc aacaatgaga acccgatatt tcttcctttg gggtgtgatt gtaacttaaa 1228
aggagacttt ttgtttttat tcttagattt gtagctctct gcatagtgag cataaattga 1288
tgtaatatgg ttttaagagat tcggtgttct ctggtgtgtg ctgcaaccac ataattgggtg 1348
atagataggt ttagttatat aagcaaattg attagagata aggggagaca tatttgatgg 1408
tctttt 1413

```

<210> 60  
 <211> 360  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (248)...(308)  
 <223> Conserved domain

```

<400> 60
Met Gly Ser Asn Glu Glu Gly Asn Pro Thr Asn Asn Ser Asp Lys Pro
  1              5              10              15
Ser Gln Ala Ala Ala Pro Glu Gln Ser Asn Val His Val Tyr His His
      20              25              30
Asp Trp Ala Ala Met Gln Ala Tyr Tyr Gly Pro Arg Val Gly Ile Pro
      35              40              45
Gln Tyr Tyr Asn Ser Asn Leu Ala Pro Gly His Ala Pro Pro Pro Tyr
      50              55              60
Met Trp Ala Ser Pro Ser Pro Met Met Ala Pro Tyr Gly Ala Pro Tyr
      65              70              75              80
Pro Pro Phe Cys Pro Pro Gly Gly Val Tyr Ala His Pro Gly Val Gln
      85              90              95
Met Gly Ser Gln Pro Gln Gly Pro Val Ser Gln Ser Ala Ser Gly Val
      100              105              110
Thr Thr Pro Leu Thr Ile Asp Ala Pro Ala Asn Ser Ala Gly Asn Ser
      115              120              125
Asp His Gly Phe Met Lys Lys Leu Lys Glu Phe Asp Gly Leu Ala Met
      130              135              140
Ser Ile Ser Asn Asn Lys Val Gly Ser Ala Glu His Ser Ser Ser Glu
      145              150              155              160
His Arg Ser Ser Gln Ser Ser Glu Asn Asp Gly Ser Ser Asn Gly Ser
      165              170              175
Asp Gly Asn Thr Thr Gly Gly Glu Gln Ser Arg Arg Lys Arg Arg Gln
      180              185              190
Gln Arg Ser Pro Ser Thr Gly Glu Arg Pro Ser Ser Gln Asn Ser Leu
      195              200              205

```



Pro Leu Arg Gly Glu Asn Glu Lys Pro Asp Val Thr Met Gly Thr Pro  
 210 215 220  
 Val Met Pro Thr Ala Met Ser Phe Gln Asn Ser Ala Gly Met Asn Gly  
 225 230 235 240  
 Val Pro Gln Pro Trp Asn Glu Lys Glu Val Lys Arg Glu Lys Arg Lys  
 245 250 255  
 Gln Ser Asn Arg Glu Ser Ala Arg Arg Ser Arg Leu Arg Lys Gln Ala  
 260 265 270  
 Glu Thr Glu Gln Leu Ser Val Lys Val Asp Ala Leu Val Ala Glu Asn  
 275 280 285  
 Met Ser Leu Arg Ser Lys Leu Gly Gln Leu Asn Asn Glu Ser Glu Lys  
 290 295 300  
 Leu Arg Leu Glu Asn Glu Ala Ile Leu Asp Gln Leu Lys Ala Gln Ala  
 305 310 315 320  
 Thr Gly Lys Thr Glu Asn Leu Ile Ser Arg Val Asp Lys Asn Asn Ser  
 325 330 335  
 Val Ser Gly Ser Lys Thr Val Gln His Gln Leu Leu Asn Ala Ser Pro  
 340 345 350  
 Ile Thr Asp Pro Val Ala Ala Ser  
 355 360

&lt;210&gt; 61

&lt;211&gt; 480

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(480)

&lt;400&gt; 61

atg ggt ctt cct gaa gat ttc atc acc gag ctt cag att cca ggt tac 48  
 Met Gly Leu Pro Glu Asp Phe Ile Thr Glu Leu Gln Ile Pro Gly Tyr  
 1 5 10 15  
 ata tta aag ata ctt tac gtc atc ggt ttc ttt aga gac atg gtc gat 96  
 Ile Leu Lys Ile Leu Tyr Val Ile Gly Phe Phe Arg Asp Met Val Asp  
 20 25 30  
 gct ctt tgt cct tac att ggt cta cct agt ttt cta gac cac aac gag 144  
 Ala Leu Cys Pro Tyr Ile Gly Leu Pro Ser Phe Leu Asp His Asn Glu  
 35 40 45  
 acc tct gga ccc gat ccg acc cga cac gct ctc tct acg tca gcg agt 192  
 Thr Ser Gly Pro Asp Pro Thr Arg His Ala Leu Ser Thr Ser Ala Ser  
 50 55 60  
 ctt gct aac gag ttg atc ccg gtg gtt cgg ttc tcg gat ctt ccg acc 240  
 Leu Ala Asn Glu Leu Ile Pro Val Val Arg Phe Ser Asp Leu Pro Thr  
 65 70 75 80  
 gat ccg gaa gat tgt tgt acg gtt tgt ttg tca gat ttt gag tcc gac 288  
 Asp Pro Glu Asp Cys Cys Thr Val Cys Leu Ser Asp Phe Glu Ser Asp  
 85 90 95  
 gat aag gtt agg cag cta ccc aag tgt gga cac gtg ttt cat cat cat 336

```

Asp Lys Val Arg Gln Leu Pro Lys Cys Gly His Val Phe His His His
      100                      105                      110

tgt tta gac cgt tgg atc gtt gac tac aac aag atg aaa tgt ccg gtt   384
Cys Leu Asp Arg Trp Ile Val Asp Tyr Asn Lys Met Lys Cys Pro Val
      115                      120                      125

tgt cgg cac cgg ttc tta ccg aaa gaa aag tac acg caa tgt gat tgg   432
Cys Arg His Arg Phe Leu Pro Lys Glu Lys Tyr Thr Gln Cys Asp Trp
      130                      135                      140

ggg tct ggt tca gat tgg ttt agt gat gaa gtg gaa agt acc aac taa   480
Gly Ser Gly Ser Asp Trp Phe Ser Asp Glu Val Glu Ser Thr Asn *
      145                      150                      155

```

<210> 62  
 <211> 159  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (86)...(129)  
 <223> Conserved domain

```

<400> 62
Met Gly Leu Pro Glu Asp Phe Ile Thr Glu Leu Gln Ile Pro Gly Tyr
  1                      5                      10                      15
Ile Leu Lys Ile Leu Tyr Val Ile Gly Phe Phe Arg Asp Met Val Asp
      20                      25                      30
Ala Leu Cys Pro Tyr Ile Gly Leu Pro Ser Phe Leu Asp His Asn Glu
      35                      40                      45
Thr Ser Gly Pro Asp Pro Thr Arg His Ala Leu Ser Thr Ser Ala Ser
      50                      55                      60
Leu Ala Asn Glu Leu Ile Pro Val Val Arg Phe Ser Asp Leu Pro Thr
      65                      70                      75                      80
Asp Pro Glu Asp Cys Cys Thr Val Cys Leu Ser Asp Phe Glu Ser Asp
      85                      90                      95
Asp Lys Val Arg Gln Leu Pro Lys Cys Gly His Val Phe His His His
      100                      105                      110
Cys Leu Asp Arg Trp Ile Val Asp Tyr Asn Lys Met Lys Cys Pro Val
      115                      120                      125
Cys Arg His Arg Phe Leu Pro Lys Glu Lys Tyr Thr Gln Cys Asp Trp
      130                      135                      140
Gly Ser Gly Ser Asp Trp Phe Ser Asp Glu Val Glu Ser Thr Asn
      145                      150                      155

```

<210> 63  
 <211> 1116  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS

&lt;222&gt; (89) ... (934)

&lt;400&gt; 63

```

ttggaaccct agaggccttt caagcaaata atcagggttaa caattttcttg atctttcttt 60
ttagcgaatt tccagttttt ggtcaata atg gca aac cct tgg tgg acg aac 112
                               Met Ala Asn Pro Trp Trp Thr Asn
                               1           5

cag agt ggt tta gcg ggc atg gtg gac cat tcg gtc tcc tca ggc cat 160
Gln Ser Gly Leu Ala Gly Met Val Asp His Ser Val Ser Ser Gly His
    10           15           20

cac caa aac cat cac cac caa agt ctt ctt acc aaa gga gat ctt gga 208
His Gln Asn His His His Gln Ser Leu Leu Thr Lys Gly Asp Leu Gly
    25           30           35           40

ata gcc atg aat cag agc caa gac aac gac caa gac gaa gaa gat gat 256
Ile Ala Met Asn Gln Ser Gln Asp Asn Asp Gln Asp Glu Glu Asp Asp
           45           50           55

cct aga gaa gga gcc gtt gag gtg gtc aac cgt aga cca aga ggt aga 304
Pro Arg Glu Gly Ala Val Glu Val Val Asn Arg Arg Pro Arg Gly Arg
    60           65           70

cca cca gga tcc aaa aac aaa ccc aaa gct cca atc ttt gtg aca aga 352
Pro Pro Gly Ser Lys Asn Lys Pro Lys Ala Pro Ile Phe Val Thr Arg
    75           80           85

gac agc ccc aac gca ctc cgt agc cat gtc ttg gag atc tcc gac ggc 400
Asp Ser Pro Asn Ala Leu Arg Ser His Val Leu Glu Ile Ser Asp Gly
    90           95           100

agt gac gtc gcc gac aca atc gct cac ttc tca aga cgc agg caa cgc 448
Ser Asp Val Ala Asp Thr Ile Ala His Phe Ser Arg Arg Arg Gln Arg
   105           110           115           120

ggc gtt tgc gtt ctc agc ggg aca ggc tca gtc gct aac gtc acc ctc 496
Gly Val Cys Val Leu Ser Gly Thr Gly Ser Val Ala Asn Val Thr Leu
           125           130           135

cgc caa gcc gcc gca cca gga ggt gtg gtc tct ctc caa ggc agg ttt 544
Arg Gln Ala Ala Ala Pro Gly Gly Val Val Ser Leu Gln Gly Arg Phe
           140           145           150

gaa atc tta tct tta acc ggt gct ttc ctc cct gga cct tcc cca ccc 592
Glu Ile Leu Ser Leu Thr Gly Ala Phe Leu Pro Gly Pro Ser Pro Pro
           155           160           165

ggg tca acc ggt tta acg gtt tac tta gcc ggg gtc cag ggt cag gtc 640
Gly Ser Thr Gly Leu Thr Val Tyr Leu Ala Gly Val Gln Gly Gln Val
           170           175           180

gtt gga ggt agc gtt gta ggc cca ctc tta gcc ata ggg tcg gtc atg 688
Val Gly Gly Ser Val Gly Pro Leu Leu Ala Ile Gly Ser Val Met
           185           190           195           200

gtg att gct gct act ttc tct aac gct act tat gag aga ttg ccc atg 736

```

```

Val Ile Ala Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Met
                205                210                215

gaa gaa gag gaa gac ggt ggc ggc tca aga cag att cac gga ggc ggt   784
Glu Glu Glu Glu Asp Gly Gly Gly Ser Arg Gln Ile His Gly Gly Gly
                220                225                230

gac tca ccg ccc aga atc ggt agt aac ctg cct gat cta tca ggg atg   832
Asp Ser Pro Pro Arg Ile Gly Ser Asn Leu Pro Asp Leu Ser Gly Met
                235                240                245

gcc ggg cca ggc tac aat atg ccg ccg cat ctg att cca aat ggg gct   880
Ala Gly Pro Gly Tyr Asn Met Pro Pro His Leu Ile Pro Asn Gly Ala
                250                255                260

ggt cag cta ggg cac gaa cca tat aca tgg gtc cac gca aga cca cct   928
Gly Gln Leu Gly His Glu Pro Tyr Thr Trp Val His Ala Arg Pro Pro
                265                270                275                280

tac tga ctcaagtgc catttctata tataatgggc tatataaata aatatataga.   984
Tyr *

tgaatataag caagcaattt gaggtagtct attacaaagc ttttgctctg gttggaaaaa 1044
taaataagta tcaaagcttt gtttgttctt aatggaaata tagagcttgg gaaggtagaa 1104
agagacgaca tt                                                    1116

```

<210> 64  
 <211> 281  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (67)...(74)  
 <223> Conserved domain

<400> 64  
 Met Ala Asn Pro Trp Trp Thr Asn Gln Ser Gly Leu Ala Gly Met Val  
 1 5 10 15  
 Asp His Ser Val Ser Ser Gly His His Gln Asn His His His Gln Ser  
 20 25 30  
 Leu Leu Thr Lys Gly Asp Leu Gly Ile Ala Met Asn Gln Ser Gln Asp  
 35 40 45  
 Asn Asp Gln Asp Glu Glu Asp Asp Pro Arg Glu Gly Ala Val Glu Val  
 50 55 60  
 Val Asn Arg Arg Pro Arg Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro  
 65 70 75 80  
 Lys Ala Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Arg Ser  
 85 90 95  
 His Val Leu Glu Ile Ser Asp Gly Ser Asp Val Ala Asp Thr Ile Ala  
 100 105 110  
 His Phe Ser Arg Arg Arg Gln Arg Gly Val Cys Val Leu Ser Gly Thr  
 115 120 125  
 Gly Ser Val Ala Asn Val Thr Leu Arg Gln Ala Ala Ala Pro Gly Gly  
 130 135 140  
 Val Val Ser Leu Gln Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly Ala

145					150.					155					160
Phe	Leu	Pro	Gly	Pro	Ser	Pro	Pro	Gly	Ser	Thr	Gly	Leu	Thr	Val	Tyr
				165					170					175	
Leu	Ala	Gly	Val	Gln	Gly	Gln	Val	Val	Gly	Gly	Ser	Val	Val	Gly	Pro
			180					185					190		
Leu	Leu	Ala	Ile	Gly	Ser	Val	Met	Val	Ile	Ala	Ala	Thr	Phe	Ser	Asn
		195					200					205			
Ala	Thr	Tyr	Glu	Arg	Leu	Pro	Met	Glu	Glu	Glu	Glu	Asp	Gly	Gly	Gly
		210				215					220				
Ser	Arg	Gln	Ile	His	Gly	Gly	Gly	Asp	Ser	Pro	Pro	Arg	Ile	Gly	Ser
225					230					235					240
Asn	Leu	Pro	Asp	Leu	Ser	Gly	Met	Ala	Gly	Pro	Gly	Tyr	Asn	Met	Pro
				245					250					255	
Pro	His	Leu	Ile	Pro	Asn	Gly	Ala	Gly	Gln	Leu	Gly	His	Glu	Pro	Tyr
			260					265					270		
Thr	Trp	Val	His	Ala	Arg	Pro	Pro	Tyr							
		275					280								

```
<210> 65
<211> 2663
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>  
<221> CDS  
<222> (31) ... (2427)
```

```

<400> 65
aagtaagaga gcttcttaag gaagaagaag atg ggt tgt gct caa tca aag atc 54
                               Met Gly Cys Ala Gln Ser Lys Ile
                               1           5

```

gag aac gaa gaa gca gtt act cgt tgc aaa gaa cga aaa caa ttg atg 102  
Glu Asn Glu Glu Ala Val Thr Arg Cys Lys Glu Arg Lys Gln Leu Met  
10 15 20

aaa gac gcc gtc act gct cgt aac gct ttc gcc gcc gct cac tca gct 150  
Lys Asp Ala Val Thr Ala Arg Asn Ala Phe Ala Ala Ala His Ser Ala  
25 30 35 40

tac	gct	atg	gct	ctt	aaa	aac	acc	gga	gct	gct	ctt	tcc	gat	tac	tct		198
Tyr	Ala	Met	Ala	Leu	Lys	Asn	Thr	Gly	Ala	Ala	Leu	Ser	Asp	Tyr	Ser		
				45					50					55			

cac ggc gag ttt tta gtc tct aat cac tcg tct tcc tcc gca gct gca 246  
 His Gly Glu Phe Leu Val Ser Asn His Ser Ser Ser Ser Ala Ala Ala  
 60 65 70

gca atc gct tct act tct tct ctt ccc act gct ata tct cct cct ctt 294  
Ala Ile Ala Ser Thr Ser Ser Leu Pro Thr Ala Ile Ser Pro Pro Leu  
75 80 85

cct tct tcc acc gct ccg gtt tct aat tca acc gct tct tct tcc tcc 342  
Pro Ser Ser Thr Ala Pro Val Ser Asn Ser Thr Ala Ser Ser Ser Ser  
90 95 100

gct gcg gtt cct cag ccg att cct gat act ctt cct cct cct cct cct	390
Ala Ala Val Pro Gln Pro Ile Pro Asp Thr Leu Pro Pro Pro Pro Pro	
105 110 115 120	
cca cca ccg ctt cct ctt caa cgt gct gct act atg ccg gag atg aac	438
Pro Pro Pro Leu Pro Leu Gln Arg Ala Ala Thr Met Pro Glu Met Asn	
125 130 135	
ggt aga tcc ggt ggt ggt cat gct ggt agt gga ctc aac gga att gaa	486
Gly Arg Ser Gly Gly Gly His Ala Gly Ser Gly Leu Asn Gly Ile Glu	
140 145 150	
gaa gat gga gcc cta gat aac gat gat gat gac gat gat gat gat gat	534
Glu Asp Gly Ala Leu Asp Asn Asp Asp Asp Asp Asp Asp Asp Asp	
155 160 165	
gac tct gaa atg gag aat cgt gat cgt ttg att agg aaa tcg aga agc	582
Asp Ser Glu Met Glu Asn Arg Asp Arg Leu Ile Arg Lys Ser Arg Ser	
170 175 180	
cgt gga ggt agt act aga gga aat agg acg acg att gaa gat cat cat	630
Arg Gly Gly Ser Thr Arg Gly Asn Arg Thr Thr Ile Glu Asp His His	
185 190 195 200	
ctt cag gag gag aaa gct ccg cca cct ccc cct ttg gcg aat tcg cgg	678
Leu Gln Glu Glu Lys Ala Pro Pro Pro Pro Pro Leu Ala Asn Ser Arg	
205 210 215	
cca att ccg ccg cca cgt cag cat cag cat caa cat cag caa cag caa	726
Pro Ile Pro Pro Pro Arg Gln His Gln His Gln His Gln Gln Gln	
220 225 230	
caa caa cct ttc tac gat tac ttc ttc cct aat gtt gag aat atg cct	774
Gln Gln Pro Phe Tyr Asp Tyr Phe Phe Pro Asn Val Glu Asn Met Pro	
235 240 245	
gga act act tta gaa gat act cct cca caa cca caa cca caa cca aca	822
Gly Thr Thr Leu Glu Asp Thr Pro Pro Gln Pro Gln Pro Gln Pro Thr	
250 255 260	
agg cct gtg cct cct caa cca cat tca cca gtc gtt act gag gat gac	870
Arg Pro Val Pro Pro Gln Pro His Ser Pro Val Val Thr Glu Asp Asp	
265 270 275 280	
gaa gat gag gag gag gaa gag gag gaa gag gag gag gaa gag gag acg	918
Glu Asp Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Thr	
285 290 295	
gtg att gaa cgg aaa cca ctg gtg gag gaa aga ccg aag aga gta gag	966
Val Ile Glu Arg Lys Pro Leu Val Glu Glu Arg Pro Lys Arg Val Glu	
300 305 310	
gaa gtg acg att gaa ttg gaa aaa gtt act aat ttg aga ggg atg aag	1014
Glu Val Thr Ile Glu Leu Glu Lys Val Thr Asn Leu Arg Gly Met Lys	
315 320 325	
aag agt aaa ggg ata ggg att ccc gga gag agg aga gga atg cga atg	1062

Lys	Ser	Lys	Gly	Ile	Gly	Ile	Pro	Gly	Glu	Arg	Arg	Gly	Met	Arg	Met	
330						335					340					
ccg	gtg	act	gcg	acg	cat	ttg	gcg	aat	gta	ttc	att	gag	ctt	gat	gat	1110
Pro	Val	Thr	Ala	Thr	His	Leu	Ala	Asn	Val	Phe	Ile	Glu	Leu	Asp	Asp	
345					350				355					360		
aat	ttc	ttg	aaa	gct	tct	gaa	agt	gct	cat	gat	gtt	tct	aag	atg	ctt	1158
Asn	Phe	Leu	Lys	Ala	Ser	Glu	Ser	Ala	His	Asp	Val	Ser	Lys	Met	Leu	
				365				370						375		
gaa	gct	act	agg	ctc	cat	tac	cat	tct	aat	ttt	gca	gat	aac	cga	gga	1206
Glu	Ala	Thr	Arg	Leu	His	Tyr	His	Ser	Asn	Phe	Ala	Asp	Asn	Arg	Gly	
			380					385					390			
cat	att	gat	cac	tct	gct	aga	gtg	atg	cgt	gta	att	aca	tgg	aat	aga	1254
His	Ile	Asp	His	Ser	Ala	Arg	Val	Met	Arg	Val	Ile	Thr	Trp	Asn	Arg	
		395				400					405					
tca	ttt	aga	gga	ata	cca	aat	gct	gat	gat	ggg	aaa	gat	gat	gtt	gat	1302
Ser	Phe	Arg	Gly	Ile	Pro	Asn	Ala	Asp	Asp	Gly	Lys	Asp	Asp	Val	Asp	
410						415				420						
ttg	gaa	gag	aat	gaa	act	cat	gct	act	gtt	ctt	gac	aaa	ttg	cta	gca	1350
Leu	Glu	Glu	Asn	Glu	Thr	His	Ala	Thr	Val	Leu	Asp	Lys	Leu	Leu	Ala	
425					430				435					440		
tgg	gaa	aag	aag	ctc	tat	gac	gaa	gtc	aag	gct	ggc	gaa	ctc	atg	aaa	1398
Trp	Glu	Lys	Lys	Leu	Tyr	Asp	Glu	Val	Lys	Ala	Gly	Glu	Leu	Met	Lys	
				445				450						455		
atc	gag	tac	cag	aaa	aag	gtt	gct	cat	tta	aat	cgg	gtg	aag	aaa	cga	1446
Ile	Glu	Tyr	Gln	Lys	Lys	Val	Ala	His	Leu	Asn	Arg	Val	Lys	Lys	Arg	
			460					465					470			
ggt	ggc	cac	tcg	gat	tca	tta	gag	aga	gct	aaa	gca	gca	gta	agt	cat	1494
Gly	Gly	His	Ser	Asp	Ser	Leu	Glu	Arg	Ala	Lys	Ala	Ala	Val	Ser	His	
		475				480					485					
ttg	cat	aca	aga	tat	ata	gtt	gat	atg	caa	tcc	atg	gac	tcc	aca	gtt	1542
Leu	His	Thr	Arg	Tyr	Ile	Val	Asp	Met	Gln	Ser	Met	Asp	Ser	Thr	Val	
	490					495					500					
tca	gaa	atc	aat	cgt	ctt	agg	gat	gaa	caa	cta	tac	cta	aag	ctc	gtt	1590
Ser	Glu	Ile	Asn	Arg	Leu	Arg	Asp	Glu	Gln	Leu	Tyr	Leu	Lys	Leu	Val	
505					510				515					520		
cac	ctt	gtt	gag	gcg	atg	ggg	aag	atg	tgg	gaa	atg	atg	caa	ata	cat	1638
His	Leu	Val	Glu	Ala	Met	Gly	Lys	Met	Trp	Glu	Met	Met	Gln	Ile	His	
				525				530						535		
cat	caa	aga	caa	gct	gag	atc	tca	aag	gtg	ttg	aga	tct	cta	gat	gtt	1686
His	Gln	Arg	Gln	Ala	Glu	Ile	Ser	Lys	Val	Leu	Arg	Ser	Leu	Asp	Val	
			540					545					550			
tca	caa	gcg	gtg	aaa	gaa	aca	aat	gat	cat	cat	cac	gaa	cgc	acc	atc	1734
Ser	Gln	Ala	Val	Lys	Glu	Thr	Asn	Asp	His	His	His	Glu	Arg	Thr	Ile	

555	560	565	
cag ctc ttg gca gtg gtt Gln Leu Leu Ala Val Val 570	caa gaa tgg cac acg Gln Glu Trp His Thr 575	cag ttt tgc agg atg Gln Phe Cys Arg Met 580	1782
ata gat cat cag aaa gaa tac ata aaa gca ctt ggc gga tgg cta aag Ile Asp His Gln Lys Glu Tyr Ile Lys Ala Leu Gly Gly Trp Leu Lys 585 590 595 600			1830
cta aat ctc atc cct atc gaa agc aca ctc aag gag aaa gta tct tcg Leu Asn Leu Ile Pro Ile Glu Ser Thr Leu Lys Glu Lys Val Ser Ser 605 610 615			1878
cct cct cga gtt ccc aat ccc gca atc caa aaa ctc ctc cac gct tgg Pro Pro Arg Val Pro Asn Pro Ala Ile Gln Lys Leu Leu His Ala Trp 620 625 630			1926
tat gac cgt tta gac aaa atc ccc gac gaa atg gct aaa agt gcc ata Tyr Asp Arg Leu Asp Lys Ile Pro Asp Glu Met Ala Lys Ser Ala Ile 635 640 645			1974
atc aat ttc gca gcg gtt gta agc acg ata atg cag cag caa gaa gac Ile Asn Phe Ala Ala Val Val Ser Thr Ile Met Gln Gln Gln Glu Asp 650 655 660			2022
gag ata agt ctc aga aac aaa tgc gaa gag aca aga aaa gaa ttg gga Glu Ile Ser Leu Arg Asn Lys Cys Glu Glu Thr Arg Lys Glu Leu Gly 665 670 675 680			2070
aga aaa att aga cag ttt gag gat tgg tac cac aaa tac atc cag aag Arg Lys Ile Arg Gln Phe Glu Asp Trp Tyr His Lys Tyr Ile Gln Lys 685 690 695			2118
aga gga ccg gag ggg atg aat ccg gat gaa gcg gat aac gat cat aat Arg Gly Pro Glu Gly Met Asn Pro Asp Glu Ala Asp Asn Asp His Asn 700 705 710			2166
gat gag gtc gct gtg agg caa ttc aat gta gaa caa att aag aag agg Asp Glu Val Ala Val Arg Gln Phe Asn Val Glu Gln Ile Lys Lys Arg 715 720 725			2214
ttg gaa gaa gaa gaa gaa gct tac cat aga caa agc cat caa gtt aga Leu Glu Glu Glu Glu Glu Ala Tyr His Arg Gln Ser His Gln Val Arg 730 735 740			2262
gag aag tca ctg gct agt ctt cga act cgc ctc ccc gag ctt ttt cag Glu Lys Ser Leu Ala Ser Leu Arg Thr Arg Leu Pro Glu Leu Phe Gln 745 750 755 760			2310
gca atg tcc gag gtt gcg tat tca tgt tcg gat atg tat aga gct ata Ala Met Ser Glu Val Ala Tyr Ser Cys Ser Asp Met Tyr Arg Ala Ile 765 770 775			2358
acg tat gcg agt aag cgg caa agc caa agc gaa cgg cat cag aaa cct Thr Tyr Ala Ser Lys Arg Gln Ser Gln Ser Glu Arg His Gln Lys Pro 780 785 790			2406



agc cag gga cag agt tcg taa gaactaatgt aagatcagag taatgtcttc 2457  
 Ser Gln Gly Gln Ser Ser \*  
 795

ttcttctttg atcttgaata tttaagcaca cacatacata caacgtatag ctaaattcttt 2517  
 atcattgctt tcttatatta aggttttggc ttttgtaaga aggtttctta catatgagat 2577  
 tcatatagtg ttgtattctt aaggaactgt tctgttgagt aataagaaag ttgtgtattg 2637  
 aaatagagtt gcatttgta attttg 2663

<210> 66  
 <211> 798  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (425)...(500)  
 <223> Conserved domain

<400> 66  
 Met Gly Cys Ala Gln Ser Lys Ile Glu Asn Glu Glu Ala Val Thr Arg  
 1 5 10 15  
 Cys Lys Glu Arg Lys Gln Leu Met Lys Asp Ala Val Thr Ala Arg Asn  
 20 25 30  
 Ala Phe Ala Ala Ala His Ser Ala Tyr Ala Met Ala Leu Lys Asn Thr  
 35 40 45  
 Gly Ala Ala Leu Ser Asp Tyr Ser His Gly Glu Phe Leu Val Ser Asn  
 50 55 60  
 His Ser Ser Ser Ser Ala Ala Ala Ala Ile Ala Ser Thr Ser Ser Leu  
 65 70 75 80  
 Pro Thr Ala Ile Ser Pro Pro Leu Pro Ser Thr Ala Pro Val Ser  
 85 90 95  
 Asn Ser Thr Ala Ser Ser Ser Ser Ala Ala Val Pro Gln Pro Ile Pro  
 100 105 110  
 Asp Thr Leu Pro Pro Pro Pro Pro Pro Pro Leu Pro Leu Gln Arg  
 115 120 125  
 Ala Ala Thr Met Pro Glu Met Asn Gly Arg Ser Gly Gly Gly His Ala  
 130 135 140  
 Gly Ser Gly Leu Asn Gly Ile Glu Glu Asp Gly Ala Leu Asp Asn Asp  
 145 150 155 160  
 Asp Asp Asp Asp Asp Asp Asp Asp Asp Ser Glu Met Glu Asn Arg Asp  
 165 170 175  
 Arg Leu Ile Arg Lys Ser Arg Ser Arg Gly Gly Ser Thr Arg Gly Asn  
 180 185 190  
 Arg Thr Thr Ile Glu Asp His His Leu Gln Glu Glu Lys Ala Pro Pro  
 195 200 205  
 Pro Pro Pro Leu Ala Asn Ser Arg Pro Ile Pro Pro Pro Arg Gln His  
 210 215 220  
 Gln His Gln His Gln Gln Gln Gln Gln Pro Phe Tyr Asp Tyr Phe  
 225 230 235 240  
 Phe Pro Asn Val Glu Asn Met Pro Gly Thr Thr Leu Glu Asp Thr Pro  
 245 250 255  
 Pro Gln Pro Gln Pro Gln Pro Thr Arg Pro Val Pro Pro Gln Pro His  
 260 265 270  
 Ser Pro Val Val Thr Glu Asp Asp Glu Asp Glu Glu Glu Glu Glu  
 275 280 285

Glu Glu Glu Glu Glu Glu Glu Thr Val Ile Glu Arg Lys Pro Leu Val  
 290 295 300  
 Glu Glu Arg Pro Lys Arg Val Glu Glu Val Thr Ile Glu Leu Glu Lys  
 305 310 315 320  
 Val Thr Asn Leu Arg Gly Met Lys Lys Ser Lys Gly Ile Gly Ile Pro  
 325 330 335  
 Gly Glu Arg Arg Gly Met Arg Met Pro Val Thr Ala Thr His Leu Ala  
 340 345 350  
 Asn Val Phe Ile Glu Leu Asp Asp Asn Phe Leu Lys Ala Ser Glu Ser  
 355 360 365  
 Ala His Asp Val Ser Lys Met Leu Glu Ala Thr Arg Leu His Tyr His  
 370 375 380  
 Ser Asn Phe Ala Asp Asn Arg Gly His Ile Asp His Ser Ala Arg Val  
 385 390 395 400  
 Met Arg Val Ile Thr Trp Asn Arg Ser Phe Arg Gly Ile Pro Asn Ala  
 405 410 415  
 Asp Asp Gly Lys Asp Asp Val Asp Leu Glu Glu Asn Glu Thr His Ala  
 420 425 430  
 Thr Val Leu Asp Lys Leu Leu Ala Trp Glu Lys Lys Leu Tyr Asp Glu  
 435 440 445  
 Val Lys Ala Gly Glu Leu Met Lys Ile Glu Tyr Gln Lys Lys Val Ala  
 450 455 460  
 His Leu Asn Arg Val Lys Lys Arg Gly Gly His Ser Asp Ser Leu Glu  
 465 470 475 480  
 Arg Ala Lys Ala Ala Val Ser His Leu His Thr Arg Tyr Ile Val Asp  
 485 490 495  
 Met Gln Ser Met Asp Ser Thr Val Ser Glu Ile Asn Arg Leu Arg Asp  
 500 505 510  
 Glu Gln Leu Tyr Leu Lys Leu Val His Leu Val Glu Ala Met Gly Lys  
 515 520 525  
 Met Trp Glu Met Met Gln Ile His His Gln Arg Gln Ala Glu Ile Ser  
 530 535 540  
 Lys Val Leu Arg Ser Leu Asp Val Ser Gln Ala Val Lys Glu Thr Asn  
 545 550 555 560  
 Asp His His His Glu Arg Thr Ile Gln Leu Leu Ala Val Val Gln Glu  
 565 570 575  
 Trp His Thr Gln Phe Cys Arg Met Ile Asp His Gln Lys Glu Tyr Ile  
 580 585 590  
 Lys Ala Leu Gly Gly Trp Leu Lys Leu Asn Leu Ile Pro Ile Glu Ser  
 595 600 605  
 Thr Leu Lys Glu Lys Val Ser Ser Pro Pro Arg Val Pro Asn Pro Ala  
 610 615 620  
 Ile Gln Lys Leu Leu His Ala Trp Tyr Asp Arg Leu Asp Lys Ile Pro  
 625 630 635 640  
 Asp Glu Met Ala Lys Ser Ala Ile Ile Asn Phe Ala Ala Val Val Ser  
 645 650 655  
 Thr Ile Met Gln Gln Gln Glu Asp Glu Ile Ser Leu Arg Asn Lys Cys  
 660 665 670  
 Glu Glu Thr Arg Lys Glu Leu Gly Arg Lys Ile Arg Gln Phe Glu Asp  
 675 680 685  
 Trp Tyr His Lys Tyr Ile Gln Lys Arg Gly Pro Glu Gly Met Asn Pro  
 690 695 700  
 Asp Glu Ala Asp Asn Asp His Asn Asp Glu Val Ala Val Arg Gln Phe  
 705 710 715 720  
 Asn Val Glu Gln Ile Lys Lys Arg Leu Glu Glu Glu Glu Glu Ala Tyr  
 725 730 735  
 His Arg Gln Ser His Gln Val Arg Glu Lys Ser Leu Ala Ser Leu Arg

			740					745					750		
Thr	Arg	Leu	Pro	Glu	Leu	Phe	Gln	Ala	Met	Ser	Glu	Val	Ala	Tyr	Ser
		755					760					765			
Cys	Ser	Asp	Met	Tyr	Arg	Ala	Ile	Thr	Tyr	Ala	Ser	Lys	Arg	Gln	Ser
	770					775					780				
Gln	Ser	Glu	Arg	His	Gln	Lys	Pro	Ser	Gln	Gly	Gln	Ser	Ser		
785					790					795					

```
<210> 67
<211> 1371
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>  
<221> CDS  
<222> (175) ... (1296)
```

```

<400> 67
atttattaag catcaatgag agaacttcag agctggggtt gagttctgtc caataataca 60
taaccacggt atcatttttg tcctttacta tctcattaca ctcttctggt attcgcccaa 120
ttcttacagt cattactctc tatagggctc gagcgggccgc ccgggcaggt ttct atg 177
                                         Met
                                         1

```

cag atg gtt cac act tcc cgc tcc att gcc cag att ggg ttc ggt gtt 225  
Gln Met Val His Thr Ser Arg Ser Ile Ala Gln Ile Gly Phe Gly Val  
5 10 15

aag tcg caa tta gta ctc act ata ggg ctc gag cgg ccg ccc ggg cag 273  
Lys Ser Gln Leu Val Leu Thr Ile Gly Leu Glu Arg Pro Pro Gly Gln  
20 25 30

gta aaa gat caa aca atg tct aaa gaa gct gag atg tcg atc gcg gtg 321  
Val Lys Asp Gln Thr Met Ser Lys Glu Ala Glu Met Ser Ile Ala Val  
35 40 45

tcg gct ttg ttc cct ggt ttt aga ttc tct cct act gat gtt gaa ctt 369  
Ser Ala Leu Phe Pro Gly Phe Arg Phe Ser Pro Thr Asp Val Glu Leu  
50 55 60 65

atc tgc tac tat ctt cgt cgt aaa atc gat ggt gat gag aac tct gtt 417  
Ile Ser Tyr Tyr Leu Arg Arg Lys Ile Asp Gly Asp Glu Asn Ser Val  
70 75 80

gct gtg att gct gag gtc gag att tac aag ttc gag ccg tgg gac ttg 465  
Ala Val Ile Ala Glu Val Glu Ile Tyr Lys Phe Glu Pro Trp Asp Leu  
85 90 95

cca gag gaa tcg .aaa ctg aaa tcg gag aac gag tgg ttt tac ttc tgc 513  
Pro Glu Glu Ser Lys Leu Lys Ser Glu Asn Glu Trp Phe Tyr Phe Cys  
100 105 110

gcg agg ggg agg aag tac ccg cac ggg tca caa agc cgg cga gcc aca 561  
Ala Arg Gly Arg Lys Tyr Pro His Gly Ser Gln Ser Arg Arg Ala Thr  
115 120 125

cag cta gga tat tgg aaa gcg acc ggt aaa gag cgg agt gtt aaa tcc	609
Gln Leu Gly Tyr Trp Lys Ala Thr Gly Lys Glu Arg Ser Val Lys Ser	
130 135 140 145	
ggg aac caa gtt gtt gga acc aag aga acg ctt gta ttt cat atc ggt	657
Gly Asn Gln Val Val Gly Thr Lys Arg Thr Leu Val Phe His Ile Gly	
150 155 160	
cgg gct cct cgt ggc gag aga acg gag tgg att atg cat gaa tac tgc	705
Arg Ala Pro Arg Gly Glu Arg Thr Glu Trp Ile Met His Glu Tyr Cys	
165 170 175	
atc cat gga gcc cca cag gat gca tta gtg gtg tgc cgg tta aga aaa	753
Ile His Gly Ala Pro Gln Asp Ala Leu Val Val Cys Arg Leu Arg Lys	
180 185 190	
aat gct gat ttt cgg gct agt tcg acc caa aaa att gag gat ggt gtt	801
Asn Ala Asp Phe Arg Ala Ser Thr Gln Lys Ile Glu Asp Gly Val	
195 200 205	
gtg caa gac gat ggc tac gtt ggc caa aga ggt ggt ttg gac aag gag	849
Val Gln Asp Asp Gly Tyr Val Gly Gln Arg Gly Gly Leu Asp Lys Glu	
210 215 220 225	
gac aaa tcc tac tat gaa tct gag cat cag ata cca aat ggt gac atc	897
Asp Lys Ser Tyr Tyr Glu Ser Glu His Gln Ile Pro Asn Gly Asp Ile	
230 235 240	
gca gaa tca tca aat gtt gtt gag gat cag gcc gat acc gat gat gat	945
Ala Glu Ser Ser Asn Val Val Glu Asp Gln Ala Asp Thr Asp Asp Asp	
245 250 255	
tgt tac gcc gag att ctg aac gat gat ata ata aag ctc gac gaa gaa	993
Cys Tyr Ala Glu Ile Leu Asn Asp Asp Ile Ile Lys Leu Asp Glu Glu	
260 265 270	
gcg ttg aaa gct agc caa gcg ttt cga cca act aat cca act cat caa	1041
Ala Leu Lys Ala Ser Gln Ala Phe Arg Pro Thr Asn Pro Thr His Gln	
275 280 285	
gaa aca ata tca agc gag tca tcg agt aag agg tca aaa tgt ggt ata	1089
Glu Thr Ile Ser Ser Glu Ser Ser Ser Lys Arg Ser Lys Cys Gly Ile	
290 295 300 305	
aaa aaa gaa tca acg gaa aca atg aat tgt tac gct ttg ttc agg atc	1137
Lys Lys Glu Ser Thr Glu Thr Met Asn Cys Tyr Ala Leu Phe Arg Ile	
310 315 320	
aag aac gtt gcc gga acc gac tcc agc tgg aga ttc ccg aac ccg ttc	1185
Lys Asn Val Ala Gly Thr Asp Ser Ser Trp Arg Phe Pro Asn Pro Phe	
325 330 335	
aaa atc aag aaa gat gat agc cag aga ttg atg aag aat gtt ctg gcc	1233
Lys Ile Lys Lys Asp Asp Ser Gln Arg Leu Met Lys Asn Val Leu Ala	
340 345 350	
act act gtt ttc ttg gct atc tta ttt tct ttc ttt tgg act gta tta	1281

Thr Thr Val Phe Leu Ala Ile Leu Phe Ser Phe Phe Trp Thr Val Leu  
 355 360 365  
 ata gct agg aac taa agctagttac gacatacata ttatttatac ataaataaat 1336  
 Ile Ala Arg Asn \*  
 370  
 atagtatttt gtctatggca aaaaaaaaaa aaaaa 1371

<210> 68  
 <211> 373  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (30)...(177)  
 <223> Conserved domain

<400> 68  
 Met Gln Met Val His Thr Ser Arg Ser Ile Ala Gln Ile Gly Phe Gly  
 1 5 10 15  
 Val Lys Ser Gln Leu Val Leu Thr Ile Gly Leu Glu Arg Pro Pro Gly  
 20 25 30  
 Gln Val Lys Asp Gln Thr Met Ser Lys Glu Ala Glu Met Ser Ile Ala  
 35 40 45  
 Val Ser Ala Leu Phe Pro Gly Phe Arg Phe Ser Pro Thr Asp Val Glu  
 50 55 60  
 Leu Ile Ser Tyr Tyr Leu Arg Arg Lys Ile Asp Gly Asp Glu Asn Ser  
 65 70 75 80  
 Val Ala Val Ile Ala Glu Val Glu Ile Tyr Lys Phe Glu Pro Trp Asp  
 85 90 95  
 Leu Pro Glu Glu Ser Lys Leu Lys Ser Glu Asn Glu Trp Phe Tyr Phe  
 100 105 110  
 Cys Ala Arg Gly Arg Lys Tyr Pro His Gly Ser Gln Ser Arg Arg Ala  
 115 120 125  
 Thr Gln Leu Gly Tyr Trp Lys Ala Thr Gly Lys Glu Arg Ser Val Lys  
 130 135 140  
 Ser Gly Asn Gln Val Val Gly Thr Lys Arg Thr Leu Val Phe His Ile  
 145 150 155 160  
 Gly Arg Ala Pro Arg Gly Glu Arg Thr Glu Trp Ile Met His Glu Tyr  
 165 170 175  
 Cys Ile His Gly Ala Pro Gln Asp Ala Leu Val Val Cys Arg Leu Arg  
 180 185 190  
 Lys Asn Ala Asp Phe Arg Ala Ser Ser Thr Gln Lys Ile Glu Asp Gly  
 195 200 205  
 Val Val Gln Asp Asp Gly Tyr Val Gly Gln Arg Gly Gly Leu Asp Lys  
 210 215 220  
 Glu Asp Lys Ser Tyr Tyr Glu Ser Glu His Gln Ile Pro Asn Gly Asp  
 225 230 235 240  
 Ile Ala Glu Ser Ser Asn Val Val Glu Asp Gln Ala Asp Thr Asp Asp  
 245 250 255  
 Asp Cys Tyr Ala Glu Ile Leu Asn Asp Asp Ile Ile Lys Leu Asp Glu  
 260 265 270  
 Glu Ala Leu Lys Ala Ser Gln Ala Phe Arg Pro Thr Asn Pro Thr His  
 275 280 285  
 Gln Glu Thr Ile Ser Ser Glu Ser Ser Ser Lys Arg Ser Lys Cys Gly

```

      290              295              300
Ile Lys Lys Glu Ser Thr Glu Thr Met Asn Cys Tyr Ala Leu Phe Arg
305              310              315              320
Ile Lys Asn Val Ala Gly Thr Asp Ser Ser Trp Arg Phe Pro Asn Pro
      325              330              335
Phe Lys Ile Lys Lys Asp Asp Ser Gln Arg Leu Met Lys Asn Val Leu
      340              345              350
Ala Thr Thr Val Phe Leu Ala Ile Leu Phe Ser Phe Phe Trp Thr Val
      355              360              365
Leu Ile Ala Arg Asn
      370

```

```

<210> 69
<211> 2217
<212> DNA
<213> Arabidopsis thaliana

```

```

<220>
<221> CDS
<222> (429)...(2135)

```

```

<221> misc_feature
<222> 2192
<223> n = A,T,C or G

```

```

<400> 69
atcttctaaaa atagttaaac tgttggtgat tttgtggaa ctgattaatt aacacaattg 60
gagaaaaacaa attgaaactt tttgtttgt tagagagtta aggagaaatt ttatctaacc 120
aagctgtgta aatctcttgt agtttttctg ccgatataca ttttcattgt gttgagggtg 180
aacgataatc aagaacgaga gagagagaga gcaagagcaa gagatttcta ctacagaaga 240
tttattatat tgatcatttt gtgtgatcaa cccataaaaa cagagagaca tagacaagtc 300
catgtttcga tgttcgac tctcttactg tctaaacggc gaaataaaaa gtctgatggg 360
tgtcacttat tgcattgata ttagtaaatc agcttgagcc caagttaaag ctgaaacttg 420
ggtttgca atg gct ggt att gat aat aaa gct gct gta atg gga gaa tgg 470
      Met Ala Gly Ile Asp Asn Lys Ala Ala Val Met Gly Glu Trp
      1              5              10

```

```

ttc gac tgt agt act act aac cac agg aag aga tcg aaa gcg gaa ctt 518
Phe Asp Cys Ser Thr Thr Asn His Arg Lys Arg Ser Lys Ala Glu Leu
15              20              25              30

```

```

ggt aga gag ttt tct tta aat tac atc aag aat gag gat tct ttg caa 566
Gly Arg Glu Phe Ser Leu Asn Tyr Ile Lys Asn Glu Asp Ser Leu Gln
      35              40              45

```

```

acc acc ttt caa gaa agt tca cga gga gct ctt cgt gaa agg att gct 614
Thr Thr Phe Gln Glu Ser Ser Arg Gly Ala Leu Arg Glu Arg Ile Ala
      50              55              60

```

```

gcg aga tcc ggg ttt aat gca ccg tgg tta aac act gag gat att ctt 662
Ala Arg Ser Gly Phe Asn Ala Pro Trp Leu Asn Thr Glu Asp Ile Leu
      65              70              75

```

```

cag tcg aaa tct tta acc atc tct tct cct ggt ctt agt cct gca act 710
Gln Ser Lys Ser Leu Thr Ile Ser Ser Pro Gly Leu Ser Pro Ala Thr
      80              85              90

```

ctg tta gag tct cct gtt ttc ctc tca aac cct ttg cta tct cca aca	758
Leu Leu Glu Ser Pro Val Phe Leu Ser Asn Pro Leu Leu Ser Pro Thr	
95 100 105 110	
acc ggg aag ctc tca tca gta cct tct gat aag gct aaa gct gag tta	806
Thr Gly Lys Leu Ser Ser Val Pro Ser Asp Lys Ala Lys Ala Glu Leu	
115 120 125	
ttt gac gac att acc aca tcc tta gcc ttc caa acc att tca gga agt	854
Phe Asp Asp Ile Thr Thr Ser Leu Ala Phe Gln Thr Ile Ser Gly Ser	
130 135 140	
ggc ctt gat cct act aac atc gct tta gaa ccc gat gat tcc caa gac	902
Gly Leu Asp Pro Thr Asn Ile Ala Leu Glu Pro Asp Asp Ser Gln Asp	
145 150 155	
tat gaa gaa aga cag ctc ggc ggt tta gga gac tcg atg gct tgt tgt	950
Tyr Glu Glu Arg Gln Leu Gly Gly Leu Gly Asp Ser Met Ala Cys Cys	
160 165 170	
gca cct gca gat gat gga tac aac tgg aga aaa tat gga caa aag cta	998
Ala Pro Ala Asp Asp Gly Tyr Asn Trp Arg Lys Tyr Gly Gln Lys Leu	
175 180 185 190	
gtt aaa gga agt gag tat ccg cgg agc tat tac aag tgc acg cac ccg	1046
Val Lys Gly Ser Glu Tyr Pro Arg Ser Tyr Tyr Lys Cys Thr His Pro	
195 200 205	
aat tgt gag gcc aag aag aag gtt gaa cgg tct cgg gaa ggt cat att	1094
Asn Cys Glu Ala Lys Lys Lys Val Glu Arg Ser Arg Glu Gly His Ile	
210 215 220	
ata gag atc ata tac aca gga gat cat ata cac agc aaa cct cca cct	1142
Ile Glu Ile Ile Tyr Thr Gly Asp His Ile His Ser Lys Pro Pro Pro	
225 230 235	
aac cgc cgg tca ggg att gga tca tcc ggt act ggc caa gac atg caa	1190
Asn Arg Arg Ser Gly Ile Gly Ser Ser Gly Thr Gly Gln Asp Met Gln	
240 245 250	
ata gat gca acc gaa tac gaa ggt ttt gct gga acc aat gag aac ata	1238
Ile Asp Ala Thr Glu Tyr Glu Gly Phe Ala Gly Thr Asn Glu Asn Ile	
255 260 265 270	
gaa tgg aca tca cct gta tct gca gag ctc gaa tac gga agc cat tca	1286
Glu Trp Thr Ser Pro Val Ser Ala Glu Leu Glu Tyr Gly Ser His Ser	
275 280 285	
gga tca atg cag gtt caa aac ggg act cat cag ttc ggg tat ggt gat	1334
Gly Ser Met Gln Val Gln Asn Gly Thr His Gln Phe Gly Tyr Gly Asp	
290 295 300	
gca gca gct gat gcc tta tat aga gat gaa aac gaa gat gat cgc acg	1382
Ala Ala Ala Asp Ala Leu Tyr Arg Asp Glu Asn Glu Asp Asp Arg Thr	
305 310 315	

tcc cac atg agt gtt tcc ctg act tac gat gga gag gta gaa gag tcc	1430
Ser His Met Ser Val Ser Leu Thr Tyr Asp Gly Glu Val Glu Glu Ser	
320 325 330	
gaa tca aag aga agg aaa cta gaa gct tat gca aca gaa acg agt gga	1478
Glu Ser Lys Arg Arg Lys Leu Glu Ala Tyr Ala Thr Glu Thr Ser Gly	
335 340 345 350	
tca acc aga gcc agc cgt gag cca aga gtt gtg gtg cag acc aca agt	1526
Ser Thr Arg Ala Ser Arg Glu Pro Arg Val Val Val Gln Thr Thr Ser	
355 360 365	
gac att gac atc ctc gat gat ggt tat cgc tgg cgc aag tat ggg caa	1574
Asp Ile Asp Ile Leu Asp Asp Gly Tyr Arg Trp Arg Lys Tyr Gly Gln	
370 375 380	
aaa gtc gtt aaa gga aac ccg aat cca agg agc tac tat aaa tgc aca	1622
Lys Val Val Lys Gly Asn Pro Asn Pro Arg Ser Tyr Tyr Lys Cys Thr	
385 390 395	
gct aat gga tgt acc gta acg aag cat gta gag aga gcc tct gat gac	1670
Ala Asn Gly Cys Thr Val Thr Lys His Val Glu Arg Ala Ser Asp Asp	
400 405 410	
ttc aag agc gta cta aca act tat ata ggc aag cac acc cac gtt gta	1718
Phe Lys Ser Val Leu Thr Thr Tyr Ile Gly Lys His Thr His Val Val	
415 420 425 430	
cca gca gca cgc aac agc agc cac gtc ggt gca ggc agt tca ggg act	1766
Pro Ala Ala Arg Asn Ser Ser His Val Gly Ala Gly Ser Ser Gly Thr	
435 440 445	
ctc caa ggc agt tta gcg act cag acc cac aac cac aat gtg cac tat	1814
Leu Gln Gly Ser Leu Ala Thr Gln Thr His Asn His Asn Val His Tyr	
450 455 460	
cca atg cca cac agt aga tct gag gga ctg gcc aca gcc aac tca tct	1862
Pro Met Pro His Ser Arg Ser Glu Gly Leu Ala Thr Ala Asn Ser Ser	
465 470 475	
cta ttt gac ttc cag tca cac ctg agg cat cct aca ggt ttc tcc gtt	1910
Leu Phe Asp Phe Gln Ser His Leu Arg His Pro Thr Gly Phe Ser Val	
480 485 490	
tac ata ggc caa tct gag ctt tct gat ctt tca atg cct ggt cta act	1958
Tyr Ile Gly Gln Ser Glu Leu Ser Asp Leu Ser Met Pro Gly Leu Thr	
495 500 505 510	
att ggg caa gag aag ctt acc agc ctg cag gcg cct gac att ggg gat	2006
Ile Gly Gln Glu Lys Leu Thr Ser Leu Gln Ala Pro Asp Ile Gly Asp	
515 520 525	
cca act ggc cta atg ttg cag tta gca gca cag ccg aag gtg gaa cca	2054
Pro Thr Gly Leu Met Leu Gln Leu Ala Ala Gln Pro Lys Val Glu Pro	
530 535 540	
gtg tca cca caa cag gga ctt gat ttg tca gcg agc tca ttg ata tgc	2102



Val Ser Pro Gln Gln Gly Leu Asp Leu Ser Ala Ser Ser Leu Ile Cys  
 545 550 555

aga gag atg ttg agt aga tta cga cag ata tga aacaaatctc ttgttcact 2155  
 Arg Glu Met Leu Ser Arg Leu Arg Gln Ile \*  
 560 565

gattgctcaa atttttgaat aaatgaaaaa ttgaganaaa aaaaaaaaaa aaaaaaaaaa 2215  
 aa 2217

<210> 70  
 <211> 568  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (178)...(234)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (372)...(428)  
 <223> Conserved domain

<400> 70  
 Met Ala Gly Ile Asp Asn Lys Ala Ala Val Met Gly Glu Trp Phe Asp  
 1 5 10 15  
 Cys Ser Thr Thr Asn His Arg Lys Arg Ser Lys Ala Glu Leu Gly Arg  
 20 25 30  
 Glu Phe Ser Leu Asn Tyr Ile Lys Asn Glu Asp Ser Leu Gln Thr Thr  
 35 40 45  
 Phe Gln Glu Ser Ser Arg Gly Ala Leu Arg Glu Arg Ile Ala Ala Arg  
 50 55 60  
 Ser Gly Phe Asn Ala Pro Trp Leu Asn Thr Glu Asp Ile Leu Gln Ser  
 65 70 75 80  
 Lys Ser Leu Thr Ile Ser Ser Pro Gly Leu Ser Pro Ala Thr Leu Leu  
 85 90 95  
 Glu Ser Pro Val Phe Leu Ser Asn Pro Leu Leu Ser Pro Thr Thr Gly  
 100 105 110  
 Lys Leu Ser Ser Val Pro Ser Asp Lys Ala Lys Ala Glu Leu Phe Asp  
 115 120 125  
 Asp Ile Thr Thr Ser Leu Ala Phe Gln Thr Ile Ser Gly Ser Gly Leu  
 130 135 140  
 Asp Pro Thr Asn Ile Ala Leu Glu Pro Asp Asp Ser Gln Asp Tyr Glu  
 145 150 155 160  
 Glu Arg Gln Leu Gly Gly Leu Gly Asp Ser Met Ala Cys Cys Ala Pro  
 165 170 175  
 Ala Asp Asp Gly Tyr Asn Trp Arg Lys Tyr Gly Gln Lys Leu Val Lys  
 180 185 190  
 Gly Ser Glu Tyr Pro Arg Ser Tyr Tyr Lys Cys Thr His Pro Asn Cys  
 195 200 205  
 Glu Ala Lys Lys Lys Val Glu Arg Ser Arg Glu Gly His Ile Ile Glu  
 210 215 220  
 Ile Ile Tyr Thr Gly Asp His Ile His Ser Lys Pro Pro Pro Asn Arg  
 225 230 235 240  
 Arg Ser Gly Ile Gly Ser Ser Gly Thr Gly Gln Asp Met Gln Ile Asp  
 245 250 255

Ala Thr Glu Tyr Glu Gly Phe Ala Gly Thr Asn Glu Asn Ile Glu Trp  
 260 265 270  
 Thr Ser Pro Val Ser Ala Glu Leu Glu Tyr Gly Ser His Ser Gly Ser  
 275 280 285  
 Met Gln Val Gln Asn Gly Thr His Gln Phe Gly Tyr Gly Asp Ala Ala  
 290 295 300  
 Ala Asp Ala Leu Tyr Arg Asp Glu Asn Glu Asp Asp Arg Thr Ser His  
 305 310 315 320  
 Met Ser Val Ser Leu Thr Tyr Asp Gly Glu Val Glu Glu Ser Glu Ser  
 325 330 335  
 Lys Arg Arg Lys Leu Glu Ala Tyr Ala Thr Glu Thr Ser Gly Ser Thr  
 340 345 350  
 Arg Ala Ser Arg Glu Pro Arg Val Val Val Gln Thr Thr Ser Asp Ile  
 355 360 365  
 Asp Ile Leu Asp Asp Gly Tyr Arg Trp Arg Lys Tyr Gly Gln Lys Val  
 370 375 380  
 Val Lys Gly Asn Pro Asn Pro Arg Ser Tyr Tyr Lys Cys Thr Ala Asn  
 385 390 395 400  
 Gly Cys Thr Val Thr Lys His Val Glu Arg Ala Ser Asp Asp Phe Lys  
 405 410 415  
 Ser Val Leu Thr Thr Tyr Ile Gly Lys His Thr His Val Val Pro Ala  
 420 425 430  
 Ala Arg Asn Ser Ser His Val Gly Ala Gly Ser Ser Gly Thr Leu Gln  
 435 440 445  
 Gly Ser Leu Ala Thr Gln Thr His Asn His Asn Val His Tyr Pro Met  
 450 455 460  
 Pro His Ser Arg Ser Glu Gly Leu Ala Thr Ala Asn Ser Ser Leu Phe  
 465 470 475 480  
 Asp Phe Gln Ser His Leu Arg His Pro Thr Gly Phe Ser Val Tyr Ile  
 485 490 495  
 Gly Gln Ser Glu Leu Ser Asp Leu Ser Met Pro Gly Leu Thr Ile Gly  
 500 505 510  
 Gln Glu Lys Leu Thr Ser Leu Gln Ala Pro Asp Ile Gly Asp Pro Thr  
 515 520 525  
 Gly Leu Met Leu Gln Leu Ala Ala Gln Pro Lys Val Glu Pro Val Ser  
 530 535 540  
 Pro Gln Gln Gly Leu Asp Leu Ser Ala Ser Ser Leu Ile Cys Arg Glu  
 545 550 555 560  
 Met Leu Ser Arg Leu Arg Gln Ile  
 565

&lt;210&gt; 71

&lt;211&gt; 1927

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (55)...(1857)

&lt;400&gt; 71

catctgatct gctctcgaag acgaaagctt cgagtactgg ttgaagctaa agct atg 57  
 Met  
 1

gga cac gtg aat cta cct gca tca aag cgt ggt aac cct cgt caa tgg 105

Gly His Val Asn Leu Pro Ala Ser Lys Arg Gly Asn Pro Arg Gln Trp	
5 10 15	
cgt ctc ctc gac atc gta acc gct gct ttc ttc ggt atc gta ctt ctc	153
Arg Leu Leu Asp Ile Val Thr Ala Ala Phe Phe Gly Ile Val Leu Leu	
20 25 30	
ttc ttc atc ctt tta ttc act cct ctt ggt gat tcc atg gcg gct tct	201
Phe Phe Ile Leu Leu Phe Thr Pro Leu Gly Asp Ser Met Ala Ala Ser	
35 40 45	
ggt cgg caa acg ctg ctt ctc tct acg gcg tca gat ccg agg caa cgg	249
Gly Arg Gln Thr Leu Leu Ser Thr Ala Ser Asp Pro Arg Gln Arg	
50 55 60 65	
cag cga tta gtg act ttg gtt gaa gct ggt cag cat ttg caa ccg atc	297
Gln Arg Leu Val Thr Leu Val Glu Ala Gly Gln His Leu Gln Pro Ile	
70 75 80	
gag tat tgt cct gcg gaa gct gtt gct cat atg cct tgt gag gat ccg	345
Glu Tyr Cys Pro Ala Glu Ala Val Arg His Met Pro Cys Glu Asp Pro	
85 90 95	
aga agg aat agt cag ctt agt aga gag atg aat ttc tat agg gag aga	393
Arg Arg Asn Ser Gln Leu Ser Arg Glu Met Asn Phe Tyr Arg Glu Arg	
100 105 110	
cat tgt cct ttg cct gag gag act ccg ctc tgt ttg att cct ccg cct	441
His Cys Pro Leu Pro Glu Glu Thr Pro Leu Cys Leu Ile Pro Pro Pro	
115 120 125	
tct ggt tat aaa att cct gtt ccg tgg cct gag agt ctt cac aag att	489
Ser Gly Tyr Lys Ile Pro Val Pro Trp Pro Glu Ser Leu His Lys Ile	
130 135 140 145	
tgg cat gca aac atg cca tat aac aaa att gct gac cgg aaa ggt cat	537
Trp His Ala Asn Met Pro Tyr Asn Lys Ile Ala Asp Arg Lys Gly His	
150 155 160	
caa gga tgg atg aaa agg gaa ggg gaa tac ttt act ttc cca ggc ggt	585
Gln Gly Trp Met Lys Arg Glu Gly Glu Tyr Phe Thr Phe Pro Gly Gly	
165 170 175	
ggc acg atg ttt cct ggc gga gct ggc caa tac att gaa aag ctt gca	633
Gly Thr Met Phe Pro Gly Gly Ala Gly Gln Tyr Ile Glu Lys Leu Ala	
180 185 190	
cag tat att ccg ctt aat ggt gga act ttg aga act gct ctt gac atg	681
Gln Tyr Ile Pro Leu Asn Gly Gly Thr Leu Arg Thr Ala Leu Asp Met	
195 200 205	
gga tgc ggg gta gct agt ttt gga ggt act cta cta tct caa ggc att	729
Gly Cys Gly Val Ala Ser Phe Gly Gly Thr Leu Leu Ser Gln Gly Ile	
210 215 220 225	
cta gcc ctc tca ttt gct cca aga gat tca cat aaa tct caa att cag	777
Leu Ala Leu Ser Phe Ala Pro Arg Asp Ser His Lys Ser Gln Ile Gln	

	230	235	240	
ttc gct ttg gaa aga gga gtg cct gca ttt gtt gcc atg ctt ggc act				825
Phe Ala Leu Glu Arg Gly Val Pro Ala Phe Val Ala Met Leu Gly Thr				
	245	250	255	
cgt aga ctc ccc ttt cct gca tac tcc ttt gac ctg atg cac tgt tcc				873
Arg Arg Leu Pro Phe Pro Ala Tyr Ser Phe Asp Leu Met His Cys Ser				
	260	265	270	
cga tgt ttg att cct ttt acg gct tac aat gca act tac ttc atc gaa				921
Arg Cys Leu Ile Pro Phe Thr Ala Tyr Asn Ala Thr Tyr Phe Ile Glu				
	275	280	285	
gta gat agg tta ctg cgc cct gga gga tat ctt gta atc tct ggc cca				969
Val Asp Arg Leu Leu Arg Pro Gly Gly Tyr Leu Val Ile Ser Gly Pro				
	290	295	300	305
cct gta caa tgg cct aaa caa gac aaa gaa tgg gct gat ctt cag gcg				1017
Pro Val Gln Trp Pro Lys Gln Asp Lys Glu Trp Ala Asp Leu Gln Ala				
	310	315	320	
gtg gct aga gct ttg tgc tat gag cta att gcg gtt gat gga aac act				1065
Val Ala Arg Ala Leu Cys Tyr Glu Leu Ile Ala Val Asp Gly Asn Thr				
	325	330	335	
gtc atc tgg aag aag cct gtt gga gat tca tgt cta cct agc cag aat				1113
Val Ile Trp Lys Lys Pro Val Gly Asp Ser Cys Leu Pro Ser Gln Asn				
	340	345	350	
gag ttt ggg ctt gag ttg tgt gat gag tct gtt ccg cca agt gat gca				1161
Glu Phe Gly Leu Glu Leu Cys Asp Glu Ser Val Pro Pro Ser Asp Ala				
	355	360	365	
tgg tat ttt aaa ttg aag agg tgt gtt acc agg cca tca tcc gtc aaa				1209
Trp Tyr Phe Lys Leu Lys Arg Cys Val Thr Arg Pro Ser Ser Val Lys				
	370	375	380	385
gga gaa cac gct ttg gga act ata tcc aag tgg ccg gag agg ctt act				1257
Gly Glu His Ala Leu Gly Thr Ile Ser Lys Trp Pro Glu Arg Leu Thr				
	390	395	400	
aaa gtt cct tct agg gcc att gtc atg aaa aac gga ttg gat gtg ttt				1305
Lys Val Pro Ser Arg Ala Ile Val Met Lys Asn Gly Leu Asp Val Phe				
	405	410	415	
gaa gca gat gca agg cgg tgg gca aga cgc gtt gct tat tac agg gat				1353
Glu Ala Asp Ala Arg Arg Trp Ala Arg Arg Val Ala Tyr Tyr Arg Asp				
	420	425	430	
tct ctt aac ttg aag ctg aaa tct cca act gtc cgc aat gtc atg gac				1401
Ser Leu Asn Leu Lys Leu Lys Ser Pro Thr Val Arg Asn Val Met Asp				
	435	440	445	
atg aac gca ttc ttc gga ggc ttt gca gca acc ctt gca tct gat cct				1449
Met Asn Ala Phe Phe Gly Gly Phe Ala Ala Thr Leu Ala Ser Asp Pro				
	450	455	460	465

gtg tgg gtt atg aat gtc att cca gct cgg aag cca tta act ctt gac 1497  
 Val Trp Val Met Asn Val Ile Pro Ala Arg Lys Pro Leu Thr Leu Asp  
 470 475 480

gtg att tat gac aga ggt ctc atc ggt gtt tac cat gat tgg tgt gaa 1545  
 Val Ile Tyr Asp Arg Gly Leu Ile Gly Val Tyr His Asp Trp Cys Glu  
 485 490 495

cca ttt tca aca tat ccc cgc acg tat gat ttc atc cat gta tca gga 1593  
 Pro Phe Ser Thr Tyr Pro Arg Thr Tyr Asp Phe Ile His Val Ser Gly  
 500 505 510

att gaa tca ctg ata aaa cga caa gac tca agc aaa tcg agg tgt agc 1641  
 Ile Glu Ser Leu Ile Lys Arg Gln Asp Ser Ser Lys Ser Arg Cys Ser  
 515 520 525

cta gta gat cta atg gta gag atg gac aga ata tta cgt cca gaa gga 1689  
 Leu Val Asp Leu Met Val Glu Met Asp Arg Ile Leu Arg Pro Glu Gly  
 530 535 540 545

aag gtt gtg atc cga gac tct cct gag gtg cta gat aaa gtc gca cga 1737  
 Lys Val Val Ile Arg Asp Ser Pro Glu Val Leu Asp Lys Val Ala Arg  
 550 555 560

atg gct cat gct gta aga tgg tct tct tcc ata cac gag aaa gaa cct 1785  
 Met Ala His Ala Val Arg Trp Ser Ser Ser Ile His Glu Lys Glu Pro  
 565 570 575

gaa tcc cat gga aga gag aag att ctt atc gca acc aaa tct ctc tgg 1833  
 Glu Ser His Gly Arg Glu Lys Ile Leu Ile Ala Thr Lys Ser Leu Trp  
 580 585 590

aaa ttg cca tca aac tcc cac tga agacacaaaa gaagaagaaa agaagaagct 1887  
 Lys Leu Pro Ser Asn Ser His \*  
 595 600

cttctcaatc ttgtaggtagc tgtcacttgc tctccagccc 1927

<210> 72  
 <211> 600  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (1)...(601)  
 <223> Conserved domain

<400> 72  
 Met Gly His Val Asn Leu Pro Ala Ser Lys Arg Gly Asn Pro Arg Gln  
 1 5 10 15  
 Trp Arg Leu Leu Asp Ile Val Thr Ala Ala Phe Phe Gly Ile Val Leu  
 20 25 30  
 Leu Phe Phe Ile Leu Leu Phe Thr Pro Leu Gly Asp Ser Met Ala Ala  
 35 40 45  
 Ser Gly Arg Gln Thr Leu Leu Leu Ser Thr Ala Ser Asp Pro Arg Gln

50	55	60
Arg Gln Arg Leu Val Thr Leu Val Glu Ala Gly Gln His Leu Gln Pro		
65	70	75
Ile Glu Tyr Cys Pro Ala Glu Ala Val Ala His Met Pro Cys Glu Asp		80
	85	90
Pro Arg Arg Asn Ser Gln Leu Ser Arg Glu Met Asn Phe Tyr Arg Glu		95
	100	105
Arg His Cys Pro Leu Pro Glu Glu Thr Pro Leu Cys Leu Ile Pro Pro		110
	115	120
Pro Ser Gly Tyr Lys Ile Pro Val Pro Trp Pro Glu Ser Leu His Lys		125
	130	135
Ile Trp His Ala Asn Met Pro Tyr Asn Lys Ile Ala Asp Arg Lys Gly		140
145	150	155
His Gln Gly Trp Met Lys Arg Glu Gly Glu Tyr Phe Thr Phe Pro Gly		160
	165	170
Gly Gly Thr Met Phe Pro Gly Gly Ala Gly Gln Tyr Ile Glu Lys Leu		175
	180	185
Ala Gln Tyr Ile Pro Leu Asn Gly Gly Thr Leu Arg Thr Ala Leu Asp		190
	195	200
Met Gly Cys Gly Val Ala Ser Phe Gly Gly Thr Leu Leu Ser Gln Gly		205
	210	215
Ile Leu Ala Leu Ser Phe Ala Pro Arg Asp Ser His Lys Ser Gln Ile		220
225	230	235
Gln Phe Ala Leu Glu Arg Gly Val Pro Ala Phe Val Ala Met Leu Gly		240
	245	250
Thr Arg Arg Leu Pro Phe Pro Ala Tyr Ser Phe Asp Leu Met His Cys		255
	260	265
Ser Arg Cys Leu Ile Pro Phe Thr Ala Tyr Asn Ala Thr Tyr Phe Ile		270
	275	280
Glu Val Asp Arg Leu Leu Arg Pro Gly Gly Tyr Leu Val Ile Ser Gly		285
	290	295
Pro Pro Val Gln Trp Pro Lys Gln Asp Lys Glu Trp Ala Asp Leu Gln		300
305	310	315
Ala Val Ala Arg Ala Leu Cys Tyr Glu Leu Ile Ala Val Asp Gly Asn		320
	325	330
Thr Val Ile Trp Lys Lys Pro Val Gly Asp Ser Cys Leu Pro Ser Gln		335
	340	345
Asn Glu Phe Gly Leu Glu Leu Cys Asp Glu Ser Val Pro Pro Ser Asp		350
	355	360
Ala Trp Tyr Phe Lys Leu Lys Arg Cys Val Thr Arg Pro Ser Ser Val		365
	370	375
Lys Gly Glu His Ala Leu Gly Thr Ile Ser Lys Trp Pro Glu Arg Leu		380
385	390	395
Thr Lys Val Pro Ser Arg Ala Ile Val Met Lys Asn Gly Leu Asp Val		400
	405	410
Phe Glu Ala Asp Ala Arg Arg Trp Ala Arg Arg Val Ala Tyr Tyr Arg		415
	420	425
Asp Ser Leu Asn Leu Lys Leu Lys Ser Pro Thr Val Arg Asn Val Met		430
	435	440
Asp Met Asn Ala Phe Phe Gly Gly Phe Ala Ala Thr Leu Ala Ser Asp		445
	450	455
Pro Val Trp Val Met Asn Val Ile Pro Ala Arg Lys Pro Leu Thr Leu		460
465	470	475
Asp Val Ile Tyr Asp Arg Gly Leu Ile Gly Val Tyr His Asp Trp Cys		480
	485	490
Glu Pro Phe Ser Thr Tyr Pro Arg Thr Tyr Asp Phe Ile His Val Ser		495
	500	505
		510

Gly	Ile	Glu	Ser	Leu	Ile	Lys	Arg	Gln	Asp	Ser	Ser	Lys	Ser	Arg	Cys
		515						520				525			
Ser	Leu	Val	Asp	Leu	Met	Val	Glu	Met	Asp	Arg	Ile	Leu	Arg	Pro	Glu
		530						535				540			
Gly	Lys	Val	Val	Ile	Arg	Asp	Ser	Pro	Glu	Val	Leu	Asp	Lys	Val	Ala
545					550					555					560
Arg	Met	Ala	His	Ala	Val	Arg	Trp	Ser	Ser	Ser	Ile	His	Glu	Lys	Glu
				565					570						575
Pro	Glu	Ser	His	Gly	Arg	Glu	Lys	Ile	Leu	Ile	Ala	Thr	Lys	Ser	Leu
			580					585					590		
Trp	Lys	Leu	Pro	Ser	Asn	Ser	His								
		595					600								

```
<210> 73
<211> 1322
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>  
<221> CDS  
<222> (51)...(1247)
```

<400> 73																	
gtc	gac	ccat	tat	cct	taaa	gta	acca	aaca	ctt	ctc	ctc	tct	tct	tct	atg	gat	56
															Met	Asp	
															1		
ttg	caa	ctg	aaa	caa	tgg	aga	agc	cag	cag	cag	caa	caa	cat	cag	aca		104
Leu	Gln	Leu	Lys	Gln	Trp	Arg	Ser	Gln	Gln	Gln	Gln	Gln	His	Gln	Thr		
		5					10						15				
gag	tca	gaa	gaa	caa	cct	tct	gca	gct	aag	ata	cca	aaa	cat	gtc	ttt		152
Glu	Ser	Glu	Glu	Gln	Pro	Ser	Ala	Ala	Lys	Ile	Pro	Lys	His	Val	Phe		
		20				25					30						
gac	cag	att	cat	tct	cac	act	gca	act	tct	act	gct	ctt	cct	ctc	ttt		200
Asp	Gln	Ile	His	Ser	His	Thr	Ala	Thr	Ser	Thr	Ala	Leu	Pro	Leu	Phe		
		35			40					45					50		
acc	cct	gag	cct	act	tct	tct	aaa	ctc	tcc	tct	ttg	tct	cct	gat	tct		248
Thr	Pro	Glu	Pro	Thr	Ser	Ser	Lys	Leu	Ser	Ser	Leu	Ser	Pro	Asp	Ser		
				55					60					65			
tcc	tcc	agg	ttc	ccc	aag	atg	ggg	agc	ttc	ttt	agc	tgg	gca	cag	tgg		296
Ser	Ser	Arg	Phe	Pro	Lys	Met	Gly	Ser	Phe	Phe	Ser	Trp	Ala	Gln	Trp		
			70				75						80				
caa	gaa	ctt	gaa	cta	caa	gct	ctg	atc	tac	agg	tac	atg	ttg	gct	ggt		344
Gln	Glu	Leu	Glu	Leu	Gln	Ala	Leu	Ile	Tyr	Arg	Tyr	Met	Leu	Ala	Gly		
		85					90					95					
gct	gct	gtt	cct	cag	gag	ctc	ctt	tta	cca	atc	aag	aaa	agc	ctt	ctc		392
Ala	Ala	Val	Pro	Gln	Glu	Leu	Leu	Leu	Pro	Ile	Lys	Lys	Ser	Leu	Leu		
		100				105					110						
cat	cta	tct	cct	tcc	tac	ttt	ctt	cac	cat	cct	ctt	caa	cac	cta	cct		440

His	Leu	Ser	Pro	Ser	Tyr	Phe	Leu	His	His	Pro	Leu	Gln	His	Leu	Pro	
115					120				125						130	
cat	tac	caa	cct	gct	tgg	tat	ttg	gga	agg	gca	gcg	atg	gat	cct	gag	488
His	Tyr	Gln	Pro	Ala	Trp	Tyr	Leu	Gly	Arg	Ala	Ala	Met	Asp	Pro	Glu	
			135					140					145			
cca	ggc	aga	tgc	agg	aga	acg	gat	ggg	aag	aag	tgg	aga	tgt	tca	aga	536
Pro	Gly	Arg	Cys	Arg	Arg	Thr	Asp	Gly	Lys	Lys	Trp	Arg	Cys	Ser	Arg	
		150						155					160			
gac	gtc	ttc	gct	ggc	cac	aag	tat	tgc	gag	cgc	cac	atg	cac	cgt	ggc	584
Asp	Val	Phe	Ala	Gly	His	Lys	Pro	Cys	Glu	Arg	His	Met	His	Arg	Gly	
	165						170					175				
cgc	aac	cgt	tca	aga	aag	cct	gtg	gaa	act	cca	acc	acc	gtc	aat	gca	632
Arg	Asn	Arg	Ser	Arg	Lys	Pro	Val	Glu	Thr	Pro	Thr	Thr	Val	Asn	Ala	
	180					185					190					
act	gcc	acg	tcc	atg	gct	tca	tca	gta	gca	gcc	gca	gcc	acc	act	aca	680
Thr	Ala	Thr	Ser	Met	Ala	Ser	Ser	Val	Ala	Ala	Ala	Ala	Thr	Thr	Thr	
195					200				205						210	
aca	gca	aca	aca	aca	tct	acg	ttt	gct	ttt	ggg	ggg	ggg	ggg	ggg	agg	728
Thr	Ala	Thr	Thr	Thr	Ser	Thr	Phe	Ala	Phe	Gly	Gly	Gly	Gly	Gly	Ser	
			215					220						225		
gag	gaa	gtg	gtt	ggg	caa	gga	gga	tct	ttc	ttc	ttc	tct	ggc	tct	tct	776
Glu	Glu	Val	Val	Gly	Gln	Gly	Gly	Ser	Phe	Phe	Phe	Ser	Gly	Ser	Ser	
		230						235					240			
aac	tct	tca	tct	gaa	ctt	ctc	cac	ctt	agt	caa	agt	tgt	tcg	gag	atg	824
Asn	Ser	Ser	Ser	Glu	Leu	Leu	His	Leu	Ser	Gln	Ser	Cys	Ser	Glu	Met	
	245						250					255				
aag	caa	gaa	agc	aac	aac	atg	aac	aac	aag	agg	cca	tac	gag	tcc	cac	872
Lys	Gln	Glu	Ser	Asn	Asn	Met	Asn	Asn	Lys	Arg	Pro	Tyr	Glu	Ser	His	
	260					265					270					
atc	gga	ttc	agt	aac	aac	aga	tca	gat	gga	gga	cac	atc	ctg	agg	ccc	920
Ile	Gly	Phe	Ser	Asn	Asn	Arg	Ser	Asp	Gly	Gly	His	Ile	Leu	Arg	Pro	
275					280				285					290		
ttc	ttt	gac	gat	tgg	cct	cgt	tct	tcg	ctc	caa	gaa	gct	gac	aat	agt	968
Phe	Phe	Asp	Asp	Trp	Pro	Arg	Ser	Ser	Leu	Gln	Glu	Ala	Asp	Asn	Ser	
			295					300					305			
tca	agc	ccc	atg	agc	tca	gcc	act	tgt	ctc	tcc	atc	tcc	atg	ccc	ggg	1016
Ser	Ser	Pro	Met	Ser	Ser	Ala	Thr	Cys	Leu	Ser	Ile	Ser	Met	Pro	Gly	
		310						315					320			
aac	tct	tcc	tca	gac	gtc	tct	ctg	aag	ctg	tcc	aca	ggc	aac	gaa	gag	1064
Asn	Ser	Ser	Ser	Asp	Val	Ser	Leu	Lys	Leu	Ser	Thr	Gly	Asn	Glu	Glu	
		325					330					335				
gga	gcc	cgg	agc	aac	aac	aat	ggg	aga	gat	cag	caa	aac	atg	agc	tgg	1112
Gly	Ala	Arg	Ser	Asn	Asn	Asn	Gly	Arg	Asp	Gln	Gln	Asn	Met	Ser	Trp	



```

      340              345              350
    tgg agc ggt gga ggt tcc aac cac cat cat cac aac atg ggc gga cca 1160
    Trp Ser Gly Gly Gly Ser Asn His His His His Asn Met Gly Gly Pro
    355              360              365              370

    ttg gcc gaa gcc ctg aga tct tct tcc tca tct tcc cca acc agt gtt 1208
    Leu Ala Glu Ala Leu Arg Ser Ser Ser Ser Ser Pro Thr Ser Val
    375              380              385

    ctc cat cag ctt ggt gtc tcg aca caa gcc ttt cat tga ccagtgtaaa 1257
    Leu His Gln Leu Gly Val Ser Thr Gln Ala Phe His *
    390              395

    accaacacaa caatgcggtt tttactgtgt ttttggtttt tatccaaatt tcctgtataa 1317
    agagg 1322

```

&lt;210&gt; 74

&lt;211&gt; 398

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 74

```

Met Asp Leu Gln Leu Lys Gln Trp Arg Ser Gln Gln Gln Gln Gln His
 1              5              10              15
Gln Thr Glu Ser Glu Glu Gln Pro Ser Ala Ala Lys Ile Pro Lys His
 20              25              30
Val Phe Asp Gln Ile His Ser His Thr Ala Thr Ser Thr Ala Leu Pro
 35              40              45
Leu Phe Thr Pro Glu Pro Thr Ser Ser Lys Leu Ser Ser Leu Ser Pro
 50              55              60
Asp Ser Ser Ser Arg Phe Pro Lys Met Gly Ser Phe Phe Ser Trp Ala
 65              70              75              80
Gln Trp Gln Glu Leu Glu Leu Gln Ala Leu Ile Tyr Arg Tyr Met Leu
 85              90              95
Ala Gly Ala Ala Val Pro Gln Glu Leu Leu Leu Pro Ile Lys Lys Ser
 100             105             110
Leu Leu His Leu Ser Pro Ser Tyr Phe Leu His His Pro Leu Gln His
 115             120             125
Leu Pro His Tyr Gln Pro Ala Trp Tyr Leu Gly Arg Ala Ala Met Asp
 130             135             140
Pro Glu Pro Gly Arg Cys Arg Arg Thr Asp Gly Lys Lys Trp Arg Cys
 145             150             155             160
Ser Arg Asp Val Phe Ala Gly His Lys Tyr Cys Glu Arg His Met His
 165             170             175
Arg Gly Arg Asn Arg Ser Arg Lys Pro Val Glu Thr Pro Thr Thr Val
 180             185             190
Asn Ala Thr Ala Thr Ser Met Ala Ser Ser Val Ala Ala Ala Thr
 195             200             205
Thr Thr Thr Ala Thr Thr Thr Thr Phe Ala Phe Gly Gly Gly Gly
 210             215             220
Gly Ser Glu Glu Val Val Gly Gln Gly Gly Ser Phe Phe Phe Ser Gly
 225             230             235             240
Ser Ser Asn Ser Ser Ser Glu Leu Leu His Leu Ser Gln Ser Cys Ser
 245             250             255
Glu Met Lys Gln Glu Ser Asn Asn Met Asn Asn Lys Arg Pro Tyr Glu
 260             265             270

```

Ser His Ile Gly Phe Ser Asn Asn Arg Ser Asp Gly Gly His Ile Leu  
 275 280 285  
 Arg Pro Phe Phe Asp Asp Trp Pro Arg Ser Ser Leu Gln Glu Ala Asp  
 290 295 300  
 Asn Ser Ser Ser Pro Met Ser Ser Ala Thr Cys Leu Ser Ile Ser Met  
 305 310 315 320  
 Pro Gly Asn Ser Ser Ser Asp Val Ser Leu Lys Leu Ser Thr Gly Asn  
 325 330 335  
 Glu Glu Gly Ala Arg Ser Asn Asn Asn Gly Arg Asp Gln Gln Asn Met  
 340 345 350  
 Ser Trp Trp Ser Gly Gly Gly Ser Asn His His His His Asn Met Gly  
 355 360 365  
 Gly Pro Leu Ala Glu Ala Leu Arg Ser Ser Ser Ser Ser Ser Pro Thr  
 370 375 380  
 Ser Val Leu His Gln Leu Gly Val Ser Thr Gln Ala Phe His  
 385 390 395

&lt;210&gt; 75

&lt;211&gt; 1191

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (50)...(1096)

&lt;400&gt; 75

ctctcaccaa cataatcaaa gaagctttcc tcacgaattc aagatcgcc atg tcc tcc 58  
 Met Ser Ser  
 1

gag gat tgg gat ctc ttc gcc gtc gtc aga agc tgc agc tct tct gtt 106  
 Glu Asp Trp Asp Leu Phe Ala Val Val Arg Ser Cys Ser Ser Ser Val  
 5 10 15

tcc acc acc aat tct tgt gct ggt cat gaa gac gac ata gga aac tgt 154  
 Ser Thr Thr Asn Ser Cys Ala Gly His Glu Asp Asp Ile Gly Asn Cys  
 20 25 30 35

aaa caa caa caa gat cct cct cct cct cct ctg ttt caa gct tct tct 202  
 Lys Gln Gln Gln Asp Pro Pro Pro Pro Pro Leu Phe Gln Ala Ser Ser  
 40 45 50

tct tgc aac gag tta caa gat tct tgc aaa cca ttt tta ccc gtt act 250  
 Ser Cys Asn Glu Leu Gln Asp Ser Cys Lys Pro Phe Leu Pro Val Thr  
 55 60 65

act act act act act act tgg tct cct cct cct cta ctt cct cct cct 298  
 Thr Thr Thr Thr Thr Thr Trp Ser Pro Pro Pro Leu Leu Pro Pro Pro  
 70 75 80

aaa gcc tca tca cca tct ccc aat atc tta cta aaa caa gaa caa gta 346  
 Lys Ala Ser Ser Pro Ser Pro Asn Ile Leu Leu Lys Gln Glu Gln Val  
 85 90 95

ctt ctc gaa tca caa gat caa aaa cct cct ctt agt gtt agg gtt ttc 394

Leu	Leu	Glu	Ser	Gln	Asp	Gln	Lys	Pro	Pro	Leu	Ser	Val	Arg	Val	Phe	
100					105					110					115	
cca	cca	tcc	act	tct	tct	tct	gtc	ttt	gtt	ttt	aga	ggt	caa	cgc	gac	442
Pro	Pro	Ser	Thr	Ser	Ser	Ser	Val	Phe	Val	Phe	Arg	Gly	Gln	Arg	Asp	
				120					125					130		
cag	ctt	ctt	caa	caa	caa	tcc	caa	cct	ccc	ctt	cga	tct	aga	aaa	aga	490
Gln	Leu	Leu	Gln	Gln	Gln	Ser	Gln	Pro	Pro	Leu	Arg	Ser	Arg	Lys	Arg	
			135					140						145		
aag	aat	cag	caa	aaa	aga	acc	ata	tgt	cat	gta	acg	caa	gag	aat	ctt	538
Lys	Asn	Gln	Gln	Lys	Arg	Thr	Ile	Cys	His	Val	Thr	Gln	Glu	Asn	Leu	
		150					155						160			
tct	tct	gat	ttg	tgg	gct	tgg	cgt	aaa	tac	ggt	caa	aaa	ccc	atc	aaa	586
Ser	Ser	Asp	Leu	Trp	Ala	Trp	Arg	Lys	Tyr	Gly	Gln	Lys	Pro	Ile	Lys	
		165				170					175					
ggc	tct	cct	tat	cca	agg	aat	tat	tac	aga	tgt	agt	agc	tca	aaa	gga	634
Gly	Ser	Pro	Tyr	Pro	Arg	Asn	Tyr	Tyr	Arg	Cys	Ser	Ser	Ser	Lys	Gly	
180					185					190					195	
tgt	tta	gca	cga	aaa	caa	gtt	gaa	aga	agt	aat	tta	gat	cct	aat	atc	682
Cys	Leu	Ala	Arg	Lys	Gln	Val	Glu	Arg	Ser	Asn	Leu	Asp	Pro	Asn	Ile	
				200					205					210		
ttc	atc	gtt	act	tac	acc	gga	gaa	cac	act	cat	cca	cgt	cct	act	cac	730
Phe	Ile	Val	Thr	Tyr	Thr	Gly	Glu	His	Thr	His	Pro	Arg	Pro	Thr	His	
			215					220					225			
cgg	aac	tct	ctc	gcc	gga	agt	act	cgt	aac	aaa	tct	cag	ccc	gtt	aac	778
Arg	Asn	Ser	Leu	Ala	Gly	Ser	Thr	Arg	Asn	Lys	Ser	Gln	Pro	Val	Asn	
		230					235					240				
ccg	gtt	cct	aaa	ccg	gac	aca	tct	cct	tta	tcg	gat	aca	gta	aaa	gaa	826
Pro	Val	Pro	Lys	Pro	Asp	Thr	Ser	Pro	Leu	Ser	Asp	Thr	Val	Lys	Glu	
		245				250					255					
gag	att	cat	ctt	tct	ccg	acg	aca	ccg	ttg	aaa	gga	aac	gat	gac	gtt	874
Glu	Ile	His	Leu	Ser	Pro	Thr	Thr	Pro	Leu	Lys	Gly	Asn	Asp	Asp	Val	
260					265				270					275		
caa	gaa	acg	aat	gga	gat	gaa	gat	atg	gtt	ggt	caa	gaa	gtc	aac	atg	922
Gln	Glu	Thr	Asn	Gly	Asp	Glu	Asp	Met	Val	Gly	Gln	Glu	Val	Asn	Met	
				280					285					290		
gaa	gag	gaa	gag	gag	gaa	gaa	gaa	gtg	gaa	gaa	gat	gat	gaa	gaa	gaa	970
Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Val	Glu	Glu	Asp	Asp	Glu	Glu	Glu	
			295					300					305			
gaa	gat	gat	gat	gac	gtg	gat	gat	ctt	ttg	ata	cca	aat	tta	gcg	gtg	1018
Glu	Asp	Asp	Asp	Asp	Val	Asp	Asp	Leu	Leu	Ile	Pro	Asn	Leu	Ala	Val	
		310					315					320				
aga	gat	cga	gat	gat	ttg	ttc	ttc	gct	gga	agt	ttt	cca	tct	tgg	tcc	1066
Arg	Asp	Arg	Asp	Asp	Leu	Phe	Phe	Ala	Gly	Ser	Phe	Pro	Ser	Trp	Ser	

<400>	76															
Met	Ser	Ser	Glu	Asp	Trp	Asp	Leu	Phe	Ala	Val	Val	Arg	Ser	Cys	Ser	
1				5					10					15		
Ser	Ser	Val	Ser	Thr	Thr	Asn	Ser	Cys	Ala	Gly	His	Glu	Asp	Asp	Ile	
			20					25					30			
Gly	Asn	Cys	Lys	Gln	Gln	Gln	Asp	Pro	Pro	Pro	Pro	Pro	Leu	Phe	Gln	
		35					40					45				
Ala	Ser	Ser	Ser	Cys	Asn	Glu	Leu	Gln	Asp	Ser	Cys	Lys	Pro	Phe	Leu	
						55					60					
Pro	Val	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Trp	Ser	Pro	Pro	Pro	Leu	Leu	
65					70					75					80	
Pro	Pro	Pro	Lys	Ala	Ser	Ser	Pro	Ser	Pro	Asn	Ile	Leu	Leu	Lys	Gln	
				85					90					95		
Glu	Gln	Val	Leu	Leu	Glu	Ser	Gln	Asp	Gln	Lys	Pro	Pro	Leu	Ser	Val	
			100					105					110			
Arg	Val	Phe	Pro	Pro	Ser	Thr	Ser	Ser	Ser	Val	Phe	Val	Phe	Arg	Gly	
		115					120					125				
Gln	Arg	Asp	Gln	Leu	Leu	Gln	Gln	Gln	Ser	Gln	Pro	Pro	Leu	Arg	Ser	
		130				135					140					
Arg	Lys	Arg	Lys	Asn	Gln	Gln	Lys	Arg	Thr	Ile	Cys	His	Val	Thr	Gln	
145				150						155					160	
Glu	Asn	Leu	Ser	Ser	Asp	Leu	Trp	Ala	Trp	Arg	Lys	Tyr	Gly	Gln	Lys	
				165					170					175		
Pro	Ile	Lys	Gly	Ser	Pro	Tyr	Pro	Arg	Asn	Tyr	Tyr	Arg	Cys	Ser	Ser	
			180					185					190			
Ser	Lys	Gly	Cys	Leu	Ala	Arg	Lys	Gln	Val	Glu	Arg	Ser	Asn	Leu	Asp	
		195					200					205				
Pro	Asn	Ile	Phe	Ile	Val	Thr	Tyr	Thr	Gly	Glu	His	Thr	His	Pro	Arg	
	210					215					220					
Pro	Thr	His	Arg	Asn	Ser	Leu	Ala	Gly	Ser	Thr	Arg	Asn	Lys	Ser	Gln	
225					230						235				240	
Pro	Val	Asn	Pro	Val	Pro	Lys	Pro	Asp	Thr	Ser	Pro	Leu	Ser	Asp	Thr	
				245					250					255		
Val	Lys	Glu	Glu	Ile	His	Leu	Ser	Pro	Thr	Thr	Pro	Leu	Lys	Gly	Asn	
			260					265					270			
Asp	Asp	Val	Gln	Glu	Thr	Asn	Gly	Asp	Glu	Asp	Met	Val	Gly	Gln	Glu	
		275					280					285				
Val	Asn	Met	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Val	Glu	Glu	Asp	Asp	

```

      290              295              300
Glu Glu Glu Glu Asp Asp Asp Asp Val Asp Asp Leu Leu Ile Pro Asn
305              310              315              320
Leu Ala Val Arg Asp Arg Asp Asp Leu Phe Phe Ala Gly Ser Phe Pro
      325              330              335
Ser Trp Ser Ala Gly Ser Ala Gly Asp Gly Gly Gly
      340              345

<210> 77
<211> 1155
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (76)...(1077)

<400> 77
attcacatta ctaatctctc aagatttcac aattttcttg tgattttctc tcagtttctt 60
atttcgtttc ataac atg gat gcc atg agt agc gta gac gag agc tct aca 111
      Met Asp Ala Met Ser Ser Val Asp Glu Ser Ser Thr
              1              5              10

act aca gat tcc att ccg gcg aga aag tca tcg tct ccg gcg agt tta 159
Thr Thr Asp Ser Ile Pro Ala Arg Lys Ser Ser Ser Pro Ala Ser Leu
      15              20              25

cta tat aga atg gga agc gga aca agc gtg gta ctt gat tca gag aac 207
Leu Tyr Arg Met Gly Ser Gly Thr Ser Val Val Leu Asp Ser Glu Asn
      30              35              40

ggt gtc gaa gtc gaa gtc gaa gcc gaa tca aga aag ctt cct tct tca 255
Gly Val Glu Val Glu Val Glu Ala Glu Ser Arg Lys Leu Pro Ser Ser
      45              50              55              60

aga ttc aaa ggt gtt gtt cct caa cca aat gga aga tgg gga gct cag 303
Arg Phe Lys Gly Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln
      65              70              75

att tac gag aaa cat caa cgc gtg tgg ctt ggt act ttc aac gag gaa 351
Ile Tyr Glu Lys His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Glu
      80              85              90

gac gaa gca gct cgt gct tac gac gtc gcg gct cac cgt ttc cgt ggc 399
Asp Glu Ala Ala Arg Ala Tyr Asp Val Ala Ala His Arg Phe Arg Gly
      95              100              105

cgc gat gcc gtt act aat ttc aaa gac acg acg ttc gaa gaa gag gtt 447
Arg Asp Ala Val Thr Asn Phe Lys Asp Thr Thr Phe Glu Glu Glu Val
      110              115              120

gag ttc tta aac gcg cat tcg aaa tca gag atc gta gat atg ttg aga 495
Glu Phe Leu Asn Ala His Ser Lys Ser Glu Ile Val Asp Met Leu Arg
      125              130              135              140

aaa cac act tac aaa gaa gag tta gac caa agg aaa cgt aac cgt gac 543

```

Lys His Thr Tyr Lys Glu Glu Leu Asp Gln Arg Lys Arg Asn Arg Asp  
 145 150 155  
 ggt aac gga aaa gag acg acg gcg ttt gct ttg gct tcg atg gtg gtt 591  
 Gly Asn Gly Lys Glu Thr Thr Ala Phe Ala Leu Ala Ser Met Val Val  
 160 165 170  
 atg acg ggg ttt aaa acg gcg gag tta ctg ttt gag aaa acg gta acg 639  
 Met Thr Gly Phe Lys Thr Ala Glu Leu Leu Phe Glu Lys Thr Val Thr  
 175 180 185  
 cca agt gac gtc ggg aaa cta aac cgt tta gtt ata cca aaa cac caa 687  
 Pro Ser Asp Val Gly Lys Leu Asn Arg Leu Val Ile Pro Lys His Gln  
 190 195 200  
 gcg gag aaa cat ttt ccg tta ccg tta ggt aat aat aac gtc tcc gtt 735  
 Ala Glu Lys His Phe Pro Leu Pro Leu Gly Asn Asn Asn Val Ser Val  
 205 210 215 220  
 aaa ggt atg ctg ttg aat ttc gaa gac gtt aac ggg aaa gtg tgg agg 783  
 Lys Gly Met Leu Leu Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg  
 225 230 235  
 ttc cgt tac tct tat tgg aat agt agt caa agt tat gtg ttg acc aaa 831  
 Phe Arg Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys  
 240 245 250  
 ggt tgg agt aga ttc gtt aaa gag aag aga ctt tgt gct ggt gat ttg 879  
 Gly Trp Ser Arg Phe Val Lys Glu Lys Arg Leu Cys Ala Gly Asp Leu  
 255 260 265  
 atc agt ttt aaa aga tcc aac gat caa gat caa aaa ttc ttt atc ggg 927  
 Ile Ser Phe Lys Arg Ser Asn Asp Gln Asp Gln Lys Phe Phe Ile Gly  
 270 275 280  
 tgg aaa tcg aaa tcc ggg ttg gat cta gag acg ggt cgg gtt atg aga 975  
 Trp Lys Ser Lys Ser Gly Leu Asp Leu Glu Thr Gly Arg Val Met Arg  
 285 290 295 300  
 ttg ttt ggg gtt gat att tct tta aac gcc gtc gtt gta gtg aag gaa 1023  
 Leu Phe Gly Val Asp Ile Ser Leu Asn Ala Val Val Val Val Lys Glu  
 305 310 315  
 aca acg gag gtg tta atg tcg tcg tta agg tgt aag aag caa cga gtt 1071  
 Thr Thr Glu Val Leu Met Ser Ser Leu Arg Cys Lys Lys Gln Arg Val  
 320 325 330  
 ttg taa taacaattta acaacttggg aaagaaaaaa aagctttttg attttaattt 1127  
 Leu \*  
 ctcttcaacg ttaatcttgc tgagatta 1155  
 <210> 78  
 <211> 333  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (59)...(124)  
 <223> Conserved domain

<400> 78  
 Met Asp Ala Met Ser Ser Val Asp Glu Ser Ser Thr Thr Thr Asp Ser  
 1 5 10 15  
 Ile Pro Ala Arg Lys Ser Ser Ser Pro Ala Ser Leu Leu Tyr Arg Met  
 20 25 30  
 Gly Ser Gly Thr Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Val  
 35 40 45  
 Glu Val Glu Ala Glu Ser Arg Lys Leu Pro Ser Ser Arg Phe Lys Gly  
 50 55 60  
 Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys  
 65 70 75 80  
 His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Glu Asp Glu Ala Ala  
 85 90 95  
 Arg Ala Tyr Asp Val Ala Ala His Arg Phe Arg Gly Arg Asp Ala Val  
 100 105 110  
 Thr Asn Phe Lys Asp Thr Thr Phe Glu Glu Glu Val Glu Phe Leu Asn  
 115 120 125  
 Ala His Ser Lys Ser Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr  
 130 135 140  
 Lys Glu Glu Leu Asp Gln Arg Lys Arg Asn Arg Asp Gly Asn Gly Lys  
 145 150 155 160  
 Glu Thr Thr Ala Phe Ala Leu Ala Ser Met Val Val Met Thr Gly Phe  
 165 170 175  
 Lys Thr Ala Glu Leu Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val  
 180 185 190  
 Gly Lys Leu Asn Arg Leu Val Ile Pro Lys His Gln Ala Glu Lys His  
 195 200 205  
 Phe Pro Leu Pro Leu Gly Asn Asn Asn Val Ser Val Lys Gly Met Leu  
 210 215 220  
 Leu Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg Phe Arg Tyr Ser  
 225 230 235 240  
 Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp Ser Arg  
 245 250 255  
 Phe Val Lys Glu Lys Arg Leu Cys Ala Gly Asp Leu Ile Ser Phe Lys  
 260 265 270  
 Arg Ser Asn Asp Gln Asp Gln Lys Phe Phe Ile Gly Trp Lys Ser Lys  
 275 280 285  
 Ser Gly Leu Asp Leu Glu Thr Gly Arg Val Met Arg Leu Phe Gly Val  
 290 295 300  
 Asp Ile Ser Leu Asn Ala Val Val Val Val Lys Glu Thr Thr Glu Val  
 305 310 315 320  
 Leu Met Ser Ser Leu Arg Cys Lys Lys Gln Arg Val Leu  
 325 330

<210> 79  
 <211> 1029  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>

&lt;221&gt; CDS

&lt;222&gt; (1)... (1029)

&lt;400&gt; 79

```

atg gag aat ccg gtg ggt tta aga ttc cgt ccg acc gac aag gag atc 48
Met Glu Asn Pro Val Gly Leu Arg Phe Arg Pro Thr Asp Lys Glu Ile
  1           5           10          15

gtc gtc gat tac ctc cga cca aaa aac tcc gat agg gac acg agt cat 96
Val Val Asp Tyr Leu Arg Pro Lys Asn Ser Asp Arg Asp Thr Ser His
          20          25          30

gtt gat cga gtc att agc aca gtc act atc cgt agc ttc gac cct tgg 144
Val Asp Arg Val Ile Ser Thr Val Thr Ile Arg Ser Phe Asp Pro Trp
          35          40          45

gag tta cct tgc cag tct agg atc aaa ctg aaa gat gag tct tgg tgt 192
Glu Leu Pro Cys Gln Ser Arg Ile Lys Leu Lys Asp Glu Ser Trp Cys
          50          55          60

ttc ttc agc cct aag gag aac aaa tat ggc aga ggt gat caa caa att 240
Phe Phe Ser Pro Lys Glu Asn Lys Tyr Gly Arg Gly Asp Gln Gln Ile
          65          70          75          80

aga aaa acg aaa tct ggt tac tgg aag att act ggc aaa cca aag cct 288
Arg Lys Thr Lys Ser Gly Tyr Trp Lys Ile Thr Gly Lys Pro Lys Pro
          85          90          95

atc ttg cgt aac cgc caa gag atc ggt gag aaa aag gtt ttg atg ttt 336
Ile Leu Arg Asn Arg Gln Glu Ile Gly Glu Lys Lys Val Leu Met Phe
          100          105          110

tac atg agt aag gaa ctt ggt gga tcc aaa tcc gac tgg gtt atg cac 384
Tyr Met Ser Lys Glu Leu Gly Gly Ser Lys Ser Asp Trp Val Met His
          115          120          125

gag tac cat gct ttc tct cct act cag atg atg atg aca tat aca ata 432
Glu Tyr His Ala Phe Ser Pro Thr Gln Met Met Met Thr Tyr Thr Ile
          130          135          140

tgt aaa gtt atg ttt aag ggt gac gtg aga gag att tct tct tct tct 480
Cys Lys Val Met Phe Lys Gly Asp Val Arg Glu Ile Ser Ser Ser Ser
          145          150          155          160

gct tct tat ggt agt gaa att gag cag agt cgt gac tct tta atc cct 528
Ala Ser Tyr Gly Ser Glu Ile Glu Gln Ser Arg Asp Ser Leu Ile Pro
          165          170          175

ctt ctt gtg aac gat tct gag gaa gaa gct caa atc gag gat gct ata 576
Leu Leu Val Asn Asp Ser Glu Glu Glu Ala Gln Ile Glu Asp Ala Ile
          180          185          190

cca ata gag gaa tgg gaa aca tgg ttg act gat gat ggt gtt gat gag 624
Pro Ile Glu Glu Trp Glu Thr Trp Leu Thr Asp Asp Gly Val Asp Glu
          195          200          205

cag gtg aat cat att atg aat atg aaa gat gat cgc aac aac cac agg 672

```



Gln Val Asn His Ile Met Asn Met Lys Asp Asp Arg Asn Asn His Arg  
 210 215 220

cct caa aag cca ttg act ggt gtc ttg att gac gat agt agt gat gat 720  
 Pro Gln Lys Pro Leu Thr Gly Val Leu Ile Asp Asp Ser Ser Asp Asp  
 225 230 235 240

gat gat gat tct gat ttg cta tct cca aca aca aat tct att gaa aat 768  
 Asp Asp Asp Ser Asp Leu Leu Ser Pro Thr Thr Asn Ser Ile Glu Asn  
 245 250 255

tcg agc act tgt gat agt ttt ggt agc tca gac caa atc aac tta gtg 816  
 Ser Ser Thr Cys Asp Ser Phe Gly Ser Ser Asp Gln Ile Asn Leu Val  
 260 265 270

tca cta act caa gag gtg agc caa gct ctg ata acc agt att gat aca 864  
 Ser Leu Thr Gln Glu Val Ser Gln Ala Leu Ile Thr Ser Ile Asp Thr  
 275 280 285

ccc gag aag att aag agt cct tat gat gat gca caa ggg act ggg gct 912  
 Pro Glu Lys Ile Lys Ser Pro Tyr Asp Asp Ala Gln Gly Thr Gly Ala  
 290 295 300

gga ggg caa aaa ttg ggt caa gag act cga gag aag aaa cga gct ggt 960  
 Gly Gly Gln Lys Leu Gly Gln Glu Thr Arg Glu Lys Lys Arg Ala Gly  
 305 310 315 320

ttc ttt cac agg atg ata caa ata ttc gtc aag aaa att cac caa tgt 1008  
 Phe Phe His Arg Met Ile Gln Ile Phe Val Lys Lys Ile His Gln Cys  
 325 330 335

tct tct atc tca aga aca taa 1029  
 Ser Ser Ile Ser Arg Thr \*  
 340

<210> 80  
 <211> 342  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (10)...(149)  
 <223> Conserved domain

<400> 80  
 Met Glu Asn Pro Val Gly Leu Arg Phe Arg Pro Thr Asp Lys Glu Ile  
 1 5 10 15  
 Val Val Asp Tyr Leu Arg Pro Lys Asn Ser Asp Arg Asp Thr Ser His  
 20 25 30  
 Val Asp Arg Val Ile Ser Thr Val Thr Ile Arg Ser Phe Asp Pro Trp  
 35 40 45  
 Glu Leu Pro Cys Gln Ser Arg Ile Lys Leu Lys Asp Glu Ser Trp Cys  
 50 55 60  
 Phe Phe Ser Pro Lys Glu Asn Lys Tyr Gly Arg Gly Asp Gln Gln Ile  
 65 70 75 80

Arg Lys Thr Lys Ser Gly Tyr Trp Lys Ile Thr Gly Lys Pro Lys Pro  
                   85                  90                  95  
 Ile Leu Arg Asn Arg Gln Glu Ile Gly Glu Lys Lys Val Leu Met Phe  
                   100                  105                  110  
 Tyr Met Ser Lys Glu Leu Gly Gly Ser Lys Ser Asp Trp Val Met His  
                   115                  120                  125  
 Glu Tyr His Ala Phe Ser Pro Thr Gln Met Met Met Thr Tyr Thr Ile  
                   130                  135                  140  
 Cys Lys Val Met Phe Lys Gly Asp Val Arg Glu Ile Ser Ser Ser Ser  
 145                  150                  155                  160  
 Ala Ser Tyr Gly Ser Glu Ile Glu Gln Ser Arg Asp Ser Leu Ile Pro  
                   165                  170                  175  
 Leu Leu Val Asn Asp Ser Glu Glu Glu Ala Gln Ile Glu Asp Ala Ile  
                   180                  185                  190  
 Pro Ile Glu Glu Trp Glu Thr Trp Leu Thr Asp Asp Gly Val Asp Glu  
                   195                  200                  205  
 Gln Val Asn His Ile Met Asn Met Lys Asp Asp Arg Asn Asn His Arg  
 210                  215                  220  
 Pro Gln Lys Pro Leu Thr Gly Val Leu Ile Asp Asp Ser Ser Asp Asp  
 225                  230                  235                  240  
 Asp Asp Asp Ser Asp Leu Leu Ser Pro Thr Thr Asn Ser Ile Glu Asn  
                   245                  250                  255  
 Ser Ser Thr Cys Asp Ser Phe Gly Ser Ser Asp Gln Ile Asn Leu Val  
                   260                  265                  270  
 Ser Leu Thr Gln Glu Val Ser Gln Ala Leu Ile Thr Ser Ile Asp Thr  
                   275                  280                  285  
 Pro Glu Lys Ile Lys Ser Pro Tyr Asp Asp Ala Gln Gly Thr Gly Ala  
                   290                  295                  300  
 Gly Gly Gln Lys Leu Gly Gln Glu Thr Arg Glu Lys Lys Arg Ala Gly  
 305                  310                  315                  320  
 Phe Phe His Arg Met Ile Gln Ile Phe Val Lys Lys Ile His Gln Cys  
                   325                  330                  335  
 Ser Ser Ile Ser Arg Thr  
                   340

<210> 81  
 <211> 1514  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (148)...(1254)

<400> 81  
 actctcttaa ctttcgtttc ttctcctacc ttcttttacc aacctttcct ttctcttaca 60  
 cacatatata tatacatata tagagagaga gaagaggaca aagagttgaa agatgaagac 120  
 tctcatgtct tcatagaaac aagtgat atg tgc gct aag aaa gaa gaa gaa gaa 174  
                   Met Cys Ala Lys Lys Glu Glu Glu Glu  
                   1                  5  
  
 gaa gaa gaa gaa gac agt tct gaa gcc atg aac aac ata caa aat tac 222  
 Glu Glu Glu Glu Asp Ser Ser Glu Ala Met Asn Asn Ile Gln Asn Tyr  
   10                  15                  20                  25  
  
 caa aat gac ctc ttc ttt cac caa ctc atc tct cat cat cac cat cat 270

Gln Asn Asp Leu Phe Phe His Gln Leu Ile Ser His His His His His	
30 35 40	
cat cat gat cct tct caa tct gaa act ttg gga gca tcc ggt aac gtt	318
His His Asp Pro Ser Gln Ser Glu Thr Leu Gly Ala Ser Gly Asn Val	
45 50 55	
gga tct ggt ttc act atc ttc tct caa gat tcc gtc tct cca ata tgg	366
Gly Ser Gly Phe Thr Ile Phe Ser Gln Asp Ser Val Ser Pro Ile Trp	
60 65 70	
tct cta cct cca cct acc tcg atc caa cca cca ttt gat cag ttt cct	414
Ser Leu Pro Pro Pro Thr Ser Ile Gln Pro Pro Phe Asp Gln Phe Pro	
75 80 85	
cct cct tct tct tct cca gca tct ttc tac gga agt ttc ttc aac aga	462
Pro Pro Ser Ser Ser Pro Ala Ser Phe Tyr Gly Ser Phe Phe Asn Arg	
90 95 100 105	
agt cga gct cat cat cag gga tta cag ttt ggg tac gag ggt ttt ggt	510
Ser Arg Ala His His Gln Gly Leu Gln Phe Gly Tyr Glu Gly Phe Gly	
110 115 120	
gga gcc acg tca gca gca cat cat cat cat gaa caa ctt cgg atc ttg	558
Gly Ala Thr Ser Ala Ala His His His His Glu Gln Leu Arg Ile Leu	
125 130 135	
tcg gaa gct tta ggt ccg gta gta caa gcc ggg tcc ggt cct ttt ggg	606
Ser Glu Ala Leu Gly Pro Val Gln Ala Gly Ser Gly Pro Phe Gly	
140 145 150	
tta caa gct gag tta ggg aag atg aca gca caa gag atc atg gac gct	654
Leu Gln Ala Glu Leu Gly Lys Met Thr Ala Gln Glu Ile Met Asp Ala	
155 160 165	
aaa gct ttg gct gct tca aag agt cat agt gaa gct gag aga aga aga	702
Lys Ala Leu Ala Ala Ser Lys Ser His Ser Glu Ala Glu Arg Arg Arg	
170 175 180 185	
aga gag aga atc aat aat cat ctc gct aag ctc cgt agc ata tta ccc	750
Arg Glu Arg Ile Asn Asn His Leu Ala Lys Leu Arg Ser Ile Leu Pro	
190 195 200	
aac acc acc aaa acg gat aaa gcg tcg tta cta gct gaa gtg atc caa	798
Asn Thr Thr Lys Thr Asp Lys Ala Ser Leu Leu Ala Glu Val Ile Gln	
205 210 215	
cat gtg aaa gag ttg aag aga gag act tca gtg atc tca gag aca aat	846
His Val Lys Glu Leu Lys Arg Glu Thr Ser Val Ile Ser Glu Thr Asn	
220 225 230	
ctt gtc cca acg gaa agc gat gag tta acg gta gct ttc acg gag gag	894
Leu Val Pro Thr Glu Ser Asp Glu Leu Thr Val Ala Phe Thr Glu Glu	
235 240 245	
gaa gaa acc gga gat ggc aga ttt gta att aaa gcg tcg ctt tgc tgt	942
Glu Glu Thr Gly Asp Gly Arg Phe Val Ile Lys Ala Ser Leu Cys Cys	

```

250          255          260          265
gaa gac agg tcg gat ctc ttg cct gac atg att aaa aca ttg aaa gct 990
Glu Asp Arg Ser Asp Leu Leu Pro Asp Met Ile Lys Thr Leu Lys Ala
          270          275          280

atg cgt ctc aaa acg ctc aag gcg gag ata acc acc gtt ggg gga cga 1038
Met Arg Leu Lys Thr Leu Lys Ala Glu Ile Thr Thr Val Gly Gly Arg
          285          290          295

gtc aag aac gtt ttg ttt gtt acc gga gaa gag agc tcc ggt gag gaa 1086
Val Lys Asn Val Leu Phe Val Thr Gly Glu Glu Ser Ser Gly Glu Glu
          300          305          310

gtg gag gaa gag tac tgt ata ggg acg att gag gaa gct ttg aaa gcg 1134
Val Glu Glu Glu Tyr Cys Ile Gly Thr Ile Glu Glu Ala Leu Lys Ala
          315          320          325

gtg atg gag aag agc aat gta gag gaa tca tct tct tct gga aat gct 1182
Val Met Glu Lys Ser Asn Val Glu Glu Ser Ser Ser Ser Gly Asn Ala
          330          335          340          345

aag aga cag aga atg agt agt cac aac act atc act atc gtc gaa caa 1230
Lys Arg Gln Arg Met Ser Ser His Asn Thr Ile Thr Ile Val Glu Gln
          350          355          360

caa caa caa tat aat cag agg taa tcaatttttt acttāaatcg cttttttttc 1284
Gln Gln Gln Tyr Asn Gln Arg *
          365

ttacttttcgg tgtatctact acgtgtgttg tttgctgggtt atggaaatga atgttgtacg 1344
tcacgtttata ctatagatat atgtgtgttt gtgtgtatgt ataacggaag tatttgtatc 1404
cgttgtggtc ttggactttt ggtttggttc taagatactt atttttaaaa acttgtatcg 1464
ttgagttggt tttctagata tgcttaatgg gagtatgtga cgaaaaaaaa 1514

<210> 82
<211> 368
<212> PRT
<213> Arabidopsis thaliana

<220>
<221> DOMAIN
<222> (167)... (242)
<223> Conserved domain

<400> 82
Met Cys Ala Lys Lys Glu Glu Glu Glu Glu Glu Glu Asp Ser Ser
 1          5          10          15
Glu Ala Met Asn Ile Gln Asn Tyr Gln Asn Asp Leu Phe Phe His
          20          25          30
Gln Leu Ile Ser His His His His His His His Asp Pro Ser Gln Ser
          35          40          45
Glu Thr Leu Gly Ala Ser Gly Asn Val Gly Ser Gly Phe Thr Ile Phe
          50          55          60
Ser Gln Asp Ser Val Ser Pro Ile Trp Ser Leu Pro Pro Pro Thr Ser
          65          70          75          80
Ile Gln Pro Pro Phe Asp Gln Phe Pro Pro Pro Ser Ser Ser Pro Ala

```

```
<210> 83
<211> 1383
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>
<221> CDS
<222> (243)...(1190)
```

```
<400> 83
ttcttgctta gtatcattct ttgtcgtggt cttttaatta accttttgca atttgtcttg 60
tgtttctcac aacacaaaaa cttgtaaaag tgtaaaaaa tcaagatctg aaaaatctta 120
tcaccgcttc taggtttttc agtttttttt cttccctttc ctgatctaaa ttaacttata 180
tttcttaggg tttcactttc tgaaacattt aatcagaatt aattaacctc tctagggtt 240
tc atg gcg aat cca tgg tgg aca gga caa gtg aac cta tcc ggc ctc 287
Met Ala Asn Pro Trp Trp Thr Gly Gln Val Asn Leu Ser Gly Leu
      1           5           10          15
```

gaa acg acg ccg cct ggt tcc tct cag tta aag aaa cca gat ctc cac 335  
Glu Thr Thr Pro Pro Gly Ser Ser Gln Leu Lys Lys Pro Asp Leu His

	20	25	30	
atc tcc atg aac atg gcc atg gac tca ggt cac aat aat cat cac cat				383
Ile Ser Met Asn Met Ala Met Asp Ser Gly His Asn Asn His His His				
	35	40	45	
cac caa gaa gtc gat aac aac aac aac gac gac gat aga gac aac ttg				431
His Gln Glu Val Asp Asn Asn Asn Asp Asp Asp Arg Asp Asn Leu				
	50	55	60	
agt gga gac gac cac gag cca cgt gaa gga gcc gta gaa gcc ccc acg				479
Ser Gly Asp Asp His Glu Pro Arg Glu Gly Ala Val Glu Ala Pro Thr				
	65	70	75	
cgc cgt cca cgt gga cgt cct gct ggt tcc aag aac aaa cca aag cca				527
Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro				
	80	85	90	95
ccg atc ttc gtc act cgc gat tct cca aat gct ctc aag agc cat gtc				575
Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Lys Ser His Val				
	100	105	110	
atg gag atc gct agt ggg act gac gtc atc gaa acc cta gct act ttt				623
Met Glu Ile Ala Ser Gly Thr Asp Val Ile Glu Thr Leu Ala Thr Phe				
	115	120	125	
gct agg cgg cgt caa cgt ggc atc tgc atc ttg agc gga aat ggc aca				671
Ala Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly Thr				
	130	135	140	
gtg gct aac gtc acc ctc cgt caa ccc tcg acc gct gcc gtt gcg gcg				719
Val Ala Asn Val Thr Leu Arg Gln Pro Ser Thr Ala Ala Val Ala Ala				
	145	150	155	
gct cct ggt ggt gcg gct gtt ttg gct tta caa ggg agg ttt gag att				767
Ala Pro Gly Gly Ala Ala Val Leu Ala Leu Gln Gly Arg Phe Glu Ile				
	160	165	170	175
ctt tct tta acc ggt tct ttc ttg cca gga ccg gct cca cct ggt tcc				815
Leu Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro Ala Pro Pro Gly Ser				
	180	185	190	
acc ggt tta acg att tac tta gcc ggt ggt caa ggt cag gtt gtt gga				863
Thr Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly				
	195	200	205	
gga agc gtg gtg ggc cca ttg atg gca gca ggt ccg gtg atg ctg atc				911
Gly Ser Val Val Gly Pro Leu Met Ala Ala Gly Pro Val Met Leu Ile				
	210	215	220	
gcc gcc acg ttc tct aac gcg act tac gag aga ttg cca ttg gag gag				959
Ala Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Leu Glu Glu				
	225	230	235	
gaa gag gca gca gag aga ggc ggt ggt gga ggc agc gga gga gtg gtt				1007
Glu Glu Ala Ala Glu Arg Gly Gly Gly Gly Ser Gly Gly Val Val				
	240	245	250	255

ccg ggg cag ctc gga ggc gga ggt tcg cca cta agc agc ggt gct ggt 1055  
 Pro Gly Gln Leu Gly Gly Gly Gly Ser Pro Leu Ser Ser Gly Ala Gly  
                   260                  265                  270

gga ggc gac ggt aac caa gga ctt ccg gtg tat aat atg ccg gga aat 1103  
 Gly Gly Asp Gly Asn Gln Gly Leu Pro Val Tyr Asn Met Pro Gly Asn  
                   275                  280                  285

ctt gtt tct aat ggt ggc agt ggt gga gga gga cag atg agc ggc caa 1151  
 Leu Val Ser Asn Gly Gly Ser Gly Gly Gly Gly Gln Met Ser Gly Gln  
                   290                  295                  300

gaa gct tat ggt tgg gct caa gct agg tca gga ttt taa cgtgcgttaa 1200  
 Glu Ala Tyr Gly Trp Ala Gln Ala Arg Ser Gly Phe \*  
                   305                  310                  315

aatggttttt aatttacaga agttaacaat aagattataa tgatgtttat tatgatgatg 1260  
 aaaaccagtc agttgctact tgttactagt gagctatata gtttgtggac attatattat 1320  
 gttctctctt gactatgatt attatttgct aaatttcact tagctaaaaa aaaaaaaaaa 1380  
 aaa 1383

<210> 84

<211> 315

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (75)...(94)

<223> Conserved domain

<221> DOMAIN

<222> (162)...(206)

<223> Conserved domain

<400> 84

Met Ala Asn Pro Trp Trp Thr Gly Gln Val Asn Leu Ser Gly Leu Glu  
   1                  5                  10                  15  
 Thr Thr Pro Pro Gly Ser Ser Gln Leu Lys Lys Pro Asp Leu His Ile  
                   20                  25                  30  
 Ser Met Asn Met Ala Met Asp Ser Gly His Asn Asn His His His His  
                   35                  40                  45  
 Gln Glu Val Asp Asn Asn Asn Asn Asp Asp Asp Arg Asp Asn Leu Ser  
                   50                  55                  60  
 Gly Asp Asp His Glu Pro Arg Glu Gly Ala Val Glu Ala Pro Thr Arg  
                   65                  70                  75                  80  
 Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro Pro  
                   85                  90                  95  
 Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Lys Ser His Val Met  
                   100                  105                  110  
 Glu Ile Ala Ser Gly Thr Asp Val Ile Glu Thr Leu Ala Thr Phe Ala  
                   115                  120                  125  
 Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly Thr Val  
                   130                  135                  140  
 Ala Asn Val Thr Leu Arg Gln Pro Ser Thr Ala Ala Val Ala Ala Ala  
                   145                  150                  155                  160

Pro	Gly	Gly	Ala	Ala	Val	Leu	Ala	Leu	Gln	Gly	Arg	Phe	Glu	Ile	Leu
				165					170					175	
Ser	Leu	Thr	Gly	Ser	Phe	Leu	Pro	Gly	Pro	Ala	Pro	Pro	Gly	Ser	Thr
			180					185					190		
Gly	Leu	Thr	Ile	Tyr	Leu	Ala	Gly	Gly	Gln	Gly	Gln	Val	Val	Gly	Gly
		195					200					205			
Ser	Val	Val	Gly	Pro	Leu	Met	Ala	Ala	Gly	Pro	Val	Met	Leu	Ile	Ala
	210					215					220				
Ala	Thr	Phe	Ser	Asn	Ala	Thr	Tyr	Glu	Arg	Leu	Pro	Leu	Glu	Glu	Glu
225					230					235					240
Glu	Ala	Ala	Glu	Arg	Gly	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Val	Val	Pro
				245					250					255	
Gly	Gln	Leu	Gly	Gly	Gly	Gly	Ser	Pro	Leu	Ser	Ser	Gly	Ala	Gly	Gly
			260					265					270		
Gly	Asp	Gly	Asn	Gln	Gly	Leu	Pro	Val	Tyr	Asn	Met	Pro	Gly	Asn	Leu
	275						280					285			
Val	Ser	Asn	Gly	Gly	Ser	Gly	Gly	Gly	Gln	Met	Ser	Gly	Gln	Glu	
	290					295				300					
Ala	Tyr	Gly	Trp	Ala	Gln	Ala	Arg	Ser	Gly	Phe					
305					310					315					

```
<210> 85
<211> 799
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>
<221> CDS
<222> (52) ... (798)
```

<400>	85																57
cgccgctcact	ctcctcccg	tgccgcacat	tagcaacact	actcccgacg	a	atg	gag									Met	Glu
															1		
acg	acg	acg	ccg	cag	tca	aaa	tca	agt	gtg	tcc	cac	cga	ccg	ccg	ttg	105	
Thr	Thr	Thr	Pro	Gln	Ser	Lys	Ser	Ser	Val	Ser	His	Arg	Pro	Pro	Leu		
			5				10				15						
gga	aga	gaa	gac	tgg	tgg	agt	gag	gaa	gcg	acg	gcg	acg	ctg	gta	gaa	153	
Gly	Arg	Glu	Asp	Trp	Trp	Ser	Glu	Glu	Ala	Thr	Ala	Thr	Leu	Val	Glu		
		20			25			30									
gcc	tgg	ggc	aat	cgt	tac	gtc	aag	ctg	aac	cac	gga	aat	ctc	cgg	cag	201	
Ala	Trp	Gly	Asn	Arg	Tyr	Val	Lys	Leu	Asn	His	Gly	Asn	Leu	Arg	Gln		
		35				40				45					50		
aat	gac	tgg	aaa	gac	gtc	gcc	gac	gcc	gtt	aac	tct	aga	cac	ggg	gat	249	
Asn	Asp	Trp	Lys	Asp	Val	Ala	Asp	Ala	Val	Asn	Ser	Arg	His	Gly	Asp		
			55				60				65						
aac	agc	cgt	aag	aag	acc	gac	tta	cag	tgt	aag	aac	cgg	gtc	gat	act	297	
Asn	Ser	Arg	Lys	Lys	Thr	Asp	Leu	Gln	Cys	Lys	Asn	Arg	Val	Asp	Thr		
			70				75				80						
ttg	aag	aag	aag	tac	aaa	aca	gag	aaa	gct	aaa	ctc	tcg	ccg	tcg	act	345	



```

Leu Lys Lys Lys Tyr Lys Thr Glu Lys Ala Lys Leu Ser Pro Ser Thr
      85                      90                      95

tgg cgt ttc tat aac cgc ctc gat gtt cta atc ggt ccc gtt gtg aag   393
Trp Arg Phe Tyr Asn Arg Leu Asp Val Leu Ile Gly Pro Val Val Lys
      100                      105                      110

aaa tcg gct ggc gga gtt gtc aaa tca gcg cct ttt aag aat cat ctg   441
Lys Ser Ala Gly Gly Val Lys Ser Ala Pro Phe Lys Asn His Leu
      115                      120                      125                      130

aat cca act gga tcg aac tct act gga agc tct ctt gaa gat gat gat   489
Asn Pro Thr Gly Ser Asn Ser Thr Gly Ser Ser Leu Glu Asp Asp Asp
      135                      140                      145

gag gat gat gat gag gtt ggt gat tgg gaa ttc gtt gct agg aag cat   537
Glu Asp Asp Asp Glu Val Gly Asp Trp Glu Phe Val Ala Arg Lys His
      150                      155                      160

cct cgt gtg gaa gag gta gat ctg agt gaa gga tca acg tgt agg gaa   585
Pro Arg Val Glu Glu Val Asp Leu Ser Glu Gly Ser Thr Cys Arg Glu
      165                      170                      175

cta gct acg gcg att ctc aag ttt gga gaa gtt tac gag aga att gaa   633
Leu Ala Thr Ala Ile Leu Lys Phe Gly Glu Val Tyr Glu Arg Ile Glu
      180                      185                      190

ggg aag aag caa cag atg atg att gag ttg gag aag cag aga atg gaa   681
Gly Lys Lys Gln Gln Met Met Ile Glu Leu Glu Lys Gln Arg Met Glu
      195                      200                      205                      210

gtg aca aag gag gta gag tta aaa cga atg aac atg ttg atg gag atg   729
Val Thr Lys Glu Val Glu Leu Lys Arg Met Asn Met Leu Met Glu Met
      215                      220                      225

cag tta gag att gag aaa tca aag cac cgg aaa cgc gca agt gct tca   777
Gln Leu Glu Ile Glu Lys Ser Lys His Arg Lys Arg Ala Ser Ala Ser
      230                      235                      240

ggt aag aag aac tca cat tag g   799
Gly Lys Lys Asn Ser His *
      245

```

&lt;210&gt; 86

&lt;211&gt; 248

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (19)...(110)

&lt;223&gt; Conserved domain

&lt;221&gt; DOMAIN

&lt;222&gt; (173)...(232)

&lt;223&gt; Conserved domain

&lt;400&gt; 86

```

Met Glu Thr Thr Thr Pro Gln Ser Lys Ser Ser Val Ser His Arg Pro
 1          5          10          15
Pro Leu Gly Arg Glu Asp Trp Trp Ser Glu Glu Ala Thr Ala Thr Leu
          20          25          30
Val Glu Ala Trp Gly Asn Arg Tyr Val Lys Leu Asn His Gly Asn Leu
          35          40          45
Arg Gln Asn Asp Trp Lys Asp Val Ala Asp Ala Val Asn Ser Arg His
          50          55          60
Gly Asp Asn Ser Arg Lys Lys Thr Asp Leu Gln Cys Lys Asn Arg Val
          65          70          75          80
Asp Thr Leu Lys Lys Lys Tyr Lys Thr Glu Lys Ala Lys Leu Ser Pro
          85          90          95
Ser Thr Trp Arg Phe Tyr Asn Arg Leu Asp Val Leu Ile Gly Pro Val
          100          105          110
Val Lys Lys Ser Ala Gly Gly Val Val Lys Ser Ala Pro Phe Lys Asn
          115          120          125
His Leu Asn Pro Thr Gly Ser Asn Ser Thr Gly Ser Ser Leu Glu Asp
          130          135          140
Asp Asp Glu Asp Asp Asp Glu Val Gly Asp Trp Glu Phe Val Ala Arg
          145          150          155          160
Lys His Pro Arg Val Glu Glu Val Asp Leu Ser Glu Gly Ser Thr Cys
          165          170          175
Arg Glu Leu Ala Thr Ala Ile Leu Lys Phe Gly Glu Val Tyr Glu Arg
          180          185          190
Ile Glu Gly Lys Lys Gln Gln Met Met Ile Glu Leu Glu Lys Gln Arg
          195          200          205
Met Glu Val Thr Lys Glu Val Glu Leu Lys Arg Met Asn Met Leu Met
          210          215          220
Glu Met Gln Leu Glu Ile Glu Lys Ser Lys His Arg Lys Arg Ala Ser
          225          230          235          240
Ala Ser Gly Lys Lys Asn Ser His
          245

```

&lt;210&gt; 87

&lt;211&gt; 866

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (46)...(837)

&lt;400&gt; 87

```

gtgtttgttag ttgaaactta ttcttcacct tttttgtttt taggt atg gag act ctg 57
                               Met Glu Thr Leu
                               1

```

```

cat cca ttc tct cac cta cct atc tct gac cac cgg ttc gtt gtt caa 105
His Pro Phe Ser His Leu Pro Ile Ser Asp His Arg Phe Val Val Gln
 5          10          15          20

```

```

gag atg gtg agc tta cac agc tcg agt agc ggt agc tgg act aaa gaa 153
Glu Met Val Ser Leu His Ser Ser Ser Ser Gly Ser Trp Thr Lys Glu
          25          30          35

```

gag aac aag atg ttc gaa cga gct ctt gcg ata tac gct gaa gac tcg	201
Glu Asn Lys Met Phe Glu Arg Ala Leu Ala Ile Tyr Ala Glu Asp Ser	
40 45 50	
cct gat cgc tgg ttt aaa gtt gct tcc atg atc cct gga aag act gtt	249
Pro Asp Arg Trp Phe Lys Val Ala Ser Met Ile Pro Gly Lys Thr Val	
55 60 65	
ttt gat gtt atg aag caa tat agt aag ctt gaa gaa gac gtt ttc gat	297
Phe Asp Val Met Lys Gln Tyr Ser Lys Leu Glu Glu Asp Val Phe Asp	
70 75 80	
att gaa gca gga cgt gtt ccc att cct ggt tat cct gca gct tct tct	345
Ile Glu Ala Gly Arg Val Pro Ile Pro Gly Tyr Pro Ala Ala Ser Ser	
85 90 95 100	
ccc ttg ggg ttt gac acg gac atg tgt cgt aaa cgg cct agt gga gct	393
Pro Leu Gly Phe Asp Thr Asp Met Cys Arg Lys Arg Pro Ser Gly Ala	
105 110 115	
aga gga tct gat caa gat cga aag aaa gga gtc cct tgg aca gag gaa	441
Arg Gly Ser Asp Gln Asp Arg Lys Lys Gly Val Pro Trp Thr Glu Glu	
120 125 130	
gaa cac agg aga ttc ttg tta ggc ctt ctc aag tac ggt aaa gga gat	489
Glu His Arg Arg Phe Leu Leu Gly Leu Leu Lys Tyr Gly Lys Gly Asp	
135 140 145	
tgg aga aac ata tcg aga aac ttc gtg gtg tca aag acg cca acg caa	537
Trp Arg Asn Ile Ser Arg Asn Phe Val Val Ser Lys Thr Pro Thr Gln	
150 155 160	
gtg gcg agc cac gcc caa aag tat tac cag aga cag ctc tcc gga gcc	585
Val Ala Ser His Ala Gln Lys Tyr Tyr Gln Arg Gln Leu Ser Gly Ala	
165 170 175 180	
aag gac aaa cgc agg cca agt atc cat gac atc aca acc ggc aat ctt	633
Lys Asp Lys Arg Arg Pro Ser Ile His Asp Ile Thr Thr Gly Asn Leu	
185 190 195	
ctc aat gcc aat ctc aac cgt tcc ttt tcc gat cat aga gat att ctc	681
Leu Asn Ala Asn Leu Asn Arg Ser Phe Ser Asp His Arg Asp Ile Leu	
200 205 210	
cct gat tta ggg ttt atc gat aag gat gat acg gag gag gga gta ata	729
Pro Asp Leu Gly Phe Ile Asp Lys Asp Asp Thr Glu Glu Gly Val Ile	
215 220 225	
ttt atg ggt cag aat ctc tct tca gaa aat ctg ttt tct cca tca cca	777
Phe Met Gly Gln Asn Leu Ser Ser Glu Asn Leu Phe Ser Pro Ser Pro	
230 235 240	
act tca ttc gaa gct gcc att aac ttc gcc gga gaa aat gtc ttc agt	825
Thr Ser Phe Glu Ala Ala Ile Asn Phe Ala Gly Glu Asn Val Phe Ser	
245 250 255 260	

gcc gga gct taa ggcaacatag aatccccaaa ctcagcggc  
Ala Gly Ala \*

866

<210> 88  
<211> 263  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<221> DOMAIN  
<222> (33)...(81)  
<223> Conserved domain

<221> DOMAIN  
<222> (129)...(183)  
<223> Conserved domain

<400> 88  
Met Glu Thr Leu His Pro Phe Ser His Leu Pro Ile Ser Asp His Arg  
1 5 10 15  
Phe Val Val Gln Glu Met Val Ser Leu His Ser Ser Ser Ser Gly Ser  
20 25 30  
Trp Thr Lys Glu Glu Asn Lys Met Phe Glu Arg Ala Leu Ala Ile Tyr  
35 40 45  
Ala Glu Asp Ser Pro Asp Arg Trp Phe Lys Val Ala Ser Met Ile Pro  
50 55 60  
Gly Lys Thr Val Phe Asp Val Met Lys Gln Tyr Ser Lys Leu Glu Glu  
65 70 75 80  
Asp Val Phe Asp Ile Glu Ala Gly Arg Val Pro Ile Pro Gly Tyr Pro  
85 90 95  
Ala Ala Ser Ser Pro Leu Gly Phe Asp Thr Asp Met Cys Arg Lys Arg  
100 105 110  
Pro Ser Gly Ala Arg Gly Ser Asp Gln Asp Arg Lys Lys Gly Val Pro  
115 120 125  
Trp Thr Glu Glu Glu His Arg Arg Phe Leu Leu Gly Leu Leu Lys Tyr  
130 135 140  
Gly Lys Gly Asp Trp Arg Asn Ile Ser Arg Asn Phe Val Val Ser Lys  
145 150 155 160  
Thr Pro Thr Gln Val Ala Ser His Ala Gln Lys Tyr Tyr Gln Arg Gln  
165 170 175  
Leu Ser Gly Ala Lys Asp Lys Arg Arg Pro Ser Ile His Asp Ile Thr  
180 185 190  
Thr Gly Asn Leu Leu Asn Ala Asn Leu Asn Arg Ser Phe Ser Asp His  
195 200 205  
Arg Asp Ile Leu Pro Asp Leu Gly Phe Ile Asp Lys Asp Asp Thr Glu  
210 215 220  
Glu Gly Val Ile Phe Met Gly Gln Asn Leu Ser Ser Glu Asn Leu Phe  
225 230 235 240  
Ser Pro Ser Pro Thr Ser Phe Glu Ala Ala Ile Asn Phe Ala Gly Glu  
245 250 255  
Asn Val Phe Ser Ala Gly Ala  
260

<210> 89

&lt;211&gt; 1200

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1200)

&lt;400&gt; 89

```

atg gaa gga gaa act cat caa tcg gag ccg tta cct tta gca tct gga 48
Met Glu Gly Glu Thr His Gln Ser Glu Pro Leu Pro Leu Ala Ser Gly
  1             5             10             15

gat tcc gat gaa ggg atc agt gct gcg att gag gcg gag ctt gct gag 96
Asp Ser Asp Glu Gly Ile Ser Ala Ala Ile Glu Ala Glu Leu Ala Glu
          20             25             30

tta gct gct gga gat tca agc ggc ggt ggt ggt tgc ggc ggc gga gga 144
Leu Ala Ala Gly Asp Ser Ser Gly Gly Gly Gly Cys Gly Gly Gly Gly
          35             40             45

gga ggg att agg agt aaa gtc aaa gga cct tgg tct acg gag gaa gac 192
Gly Gly Ile Arg Ser Lys Val Lys Gly Pro Trp Ser Thr Glu Glu Asp
          50             55             60

gcg gtt ctt acg aag ctt gtg agg aaa cta gga ccg agg aat tgg agt 240
Ala Val Leu Thr Lys Leu Val Arg Lys Leu Gly Pro Arg Asn Trp Ser
          65             70             75             80

ttg atc gct cgt gga atc cct ggt cgt tct ggt aaa tct tgt cga ttg 288
Leu Ile Ala Arg Gly Ile Pro Gly Arg Ser Gly Lys Ser Cys Arg Leu
          85             90             95

cgt tgg tgt aat cag ctt gac cct tgt ctc aag cgt aaa cct ttc tct 336
Arg Trp Cys Asn Gln Leu Asp Pro Cys Leu Lys Arg Lys Pro Phe Ser
          100             105             110

gat gag gaa gat cgt atg ata att tct gct cac gcg gtt cac gga aac 384
Asp Glu Glu Asp Arg Met Ile Ile Ser Ala His Ala Val His Gly Asn
          115             120             125

aaa tgg gca gtg att gcg aaa ctg tta act gga aga aca gat aac gct 432
Lys Trp Ala Val Ile Ala Lys Leu Leu Thr Gly Arg Thr Asp Asn Ala
          130             135             140

atc aag aac cat tgg aac tca act ctt aga cgt aaa tac gca gat cta 480
Ile Lys Asn His Trp Asn Ser Thr Leu Arg Arg Lys Tyr Ala Asp Leu
          145             150             155             160

tgg aat aat ggt caa tgg atg gct aat tca gtt act acc gct tcc gtc 528
Trp Asn Asn Gly Gln Trp Met Ala Asn Ser Val Thr Thr Ala Ser Val
          165             170             175

aaa aac gaa aat gtc gat gaa aca act aat cct cca tcc tca aaa caa 576
Lys Asn Glu Asn Val Asp Glu Thr Thr Asn Pro Pro Ser Ser Lys Gln
          180             185             190

```

cag ttg cct caa gga gat atc aat tcc tcc ccg cca aag cct cca caa 624  
 Gln Leu Pro Gln Gly Asp Ile Asn Ser Ser Pro Pro Lys Pro Pro Gln  
 195 200 205

gtg agt gat gtt gta atg gaa gag gca gct aat gaa cct caa gaa cct 672  
 Val Ser Asp Val Val Met Glu Glu Ala Ala Asn Glu Pro Gln Glu Pro  
 210 215 220

caa gaa caa caa gaa caa gct cca ccg gta gtg agt aac gtt cca aca 720  
 Gln Glu Gln Gln Glu Gln Ala Pro Pro Val Val Ser Asn Val Pro Thr  
 225 230 235 240

gaa aat aac gtg ttt cgc ccc gtg gct cgc gtt ggt gca ttc agt atc 768  
 Glu Asn Asn Val Phe Arg Pro Val Ala Arg Val Gly Ala Phe Ser Ile  
 245 250 255

tat aac cca acc agc cag aaa aat ggt tac aga gac tac aat ata gtt 816  
 Tyr Asn Pro Thr Ser Gln Lys Asn Gly Tyr Arg Asp Tyr Asn Ile Val  
 260 265 270

cca tgt gag gga ccg ttg att caa gca gct aaa ccc gat tca ttg gct 864  
 Pro Cys Glu Gly Pro Leu Ile Gln Ala Ala Lys Pro Asp Ser Leu Ala  
 275 280 285

ggt aaa ttt cta caa tct ctt tgt gat gaa cct caa atc ccg tca aaa 912  
 Gly Lys Phe Leu Gln Ser Leu Cys Asp Glu Pro Gln Ile Pro Ser Lys  
 290 295 300

tgt ggt cac ggc tgc tct act ctt cca gct gaa act aag ttc tca cgt 960  
 Cys Gly His Gly Cys Ser Thr Leu Pro Ala Glu Thr Lys Phe Ser Arg  
 305 310 315 320

aac tcg gtg ttg gga cct gag ttt gtg gat tat gaa gag cct tct gcg 1008  
 Asn Ser Val Leu Gly Pro Glu Phe Val Asp Tyr Glu Glu Pro Ser Ala  
 325 330 335

gtt ttc aac caa gaa cta ata tct ata gca acc gat ttg aac aac att 1056  
 Val Phe Asn Gln Glu Leu Ile Ser Ile Ala Thr Asp Leu Asn Asn Ile  
 340 345 350

gca tgg ata aag agt ggc ctt gac aat gct gtt gtc aga gaa gca gaa 1104  
 Ala Trp Ile Lys Ser Gly Leu Asp Asn Ala Val Val Arg Glu Ala Glu  
 355 360 365

cag agt ttg aag atg gat aat tac aat tat aat gat cct cga atc aag 1152  
 Gln Ser Leu Lys Met Asp Asn Tyr Asn Tyr Asn Asp Pro Arg Ile Lys  
 370 375 380

ttt act gga atg atg cct aga caa gat ttc ttc tgt gca aga agc tga 1200  
 Phe Thr Gly Met Met Pro Arg Gln Asp Phe Phe Cys Ala Arg Ser \*  
 385 390 395

&lt;210&gt; 90

&lt;211&gt; 399

&lt;212&gt; PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (56)...(154)

<223> Conserved domain

<400> 90

```

Met Glu Gly Glu Thr His Gln Ser Glu Pro Leu Pro Leu Ala Ser Gly
 1           5           10           15
Asp Ser Asp Glu Gly Ile Ser Ala Ala Ile Glu Ala Glu Leu Ala Glu
          20           25           30
Leu Ala Ala Gly Asp Ser Ser Gly Gly Gly Gly Cys Gly Gly Gly Gly
          35           40           45
Gly Gly Ile Arg Ser Lys Val Lys Gly Pro Trp Ser Thr Glu Glu Asp
 50           55           60
Ala Val Leu Thr Lys Leu Val Arg Lys Leu Gly Pro Arg Asn Trp Ser
65           70           75           80
Leu Ile Ala Arg Gly Ile Pro Gly Arg Ser Gly Lys Ser Cys Arg Leu
          85           90           95
Arg Trp Cys Asn Gln Leu Asp Pro Cys Leu Lys Arg Lys Pro Phe Ser
          100          105          110
Asp Glu Gly Asp Arg Met Ile Ile Ser Ala His Ala Val His Gly Asn
          115          120          125
Lys Trp Ala Val Ile Ala Lys Leu Leu Thr Gly Arg Thr Asp Asn Ala
          130          135          140
Ile Lys Asn His Trp Asn Ser Thr Leu Arg Arg Lys Tyr Ala Asp Leu
145          150          155          160
Trp Asn Asn Gly Gln Trp Met Ala Asn Ser Val Thr Thr Ala Ser Val
          165          170          175
Lys Asn Glu Asn Val Asp Glu Thr Thr Asn Pro Pro Ser Ser Lys Gln
          180          185          190
Gln Leu Pro Gln Gly Asp Ile Asn Ser Ser Pro Pro Lys Pro Pro Gln
          195          200          205
Val Ser Asp Val Val Met Glu Glu Ala Ala Asn Glu Pro Gln Glu Pro
          210          215          220
Gln Glu Gln Gln Glu Gln Ala Pro Pro Val Val Ser Asn Val Pro Thr
225          230          235          240
Glu Asn Asn Val Phe Arg Pro Val Ala Arg Val Gly Ala Phe Ser Ile
          245          250          255
Tyr Asn Pro Thr Ser Gln Lys Asn Gly Tyr Arg Asp Tyr Asn Ile Val
          260          265          270
Pro Cys Glu Gly Pro Leu Ile Gln Ala Ala Lys Pro Asp Ser Leu Ala
          275          280          285
Gly Lys Phe Leu Gln Ser Leu Cys Asp Glu Pro Gln Ile Pro Ser Lys
          290          295          300
Cys Gly His Gly Cys Ser Thr Leu Pro Ala Glu Thr Lys Phe Ser Arg
305          310          315          320
Asn Ser Val Leu Gly Pro Glu Phe Val Asp Tyr Glu Glu Pro Ser Ala
          325          330          335
Val Phe Asn Gln Glu Leu Ile Ser Ile Ala Thr Asp Leu Asn Asn Ile
          340          345          350
Ala Trp Ile Lys Ser Gly Leu Asp Asn Ala Val Val Arg Glu Ala Glu
          355          360          365
Gln Ser Leu Lys Met Asp Asn Tyr Asn Tyr Asn Asp Pro Arg Ile Lys
          370          375          380
Phe Thr Gly Met Met Pro Arg Gln Asp Phe Phe Cys Ala Arg Ser

```

395

```
<220>
<221> CDS
<222> (82) ... (879)
```

121



```

caa ggt caa gtg atc gga ggc agt gtg gtg gga ccg ctt acg gca tca 639
Gln Gly Gln Val Ile Gly Gly Ser Val Val Gly Pro Leu Thr Ala Ser
      175                      180                      185

agt ccg gtg gtc gtt atg gca gct tca ttt gga aac gca tct tac gag 687
Ser Pro Val Val Val Met Ala Ala Ser Phe Gly Asn Ala Ser Tyr Glu
      190                      195                      200

agg ctg cca cta gag gag gag gag gaa act gaa aga gaa ata gat gga 735
Arg Leu Pro Leu Glu Glu Glu Glu Glu Thr Glu Arg Glu Ile Asp Gly
      205                      210                      215

aac gcg gct agg gcg att gga acg caa acg cag aaa cag tta atg caa 783
Asn Ala Ala Arg Ala Ile Gly Thr Gln Thr Gln Lys Gln Leu Met Gln
      220                      225                      230

gat gcg aca tcg ttt att ggg tcg ccg tcg aat tta att aac tct gtt 831
Asp Ala Thr Ser Phe Ile Gly Ser Pro Ser Asn Leu Ile Asn Ser Val
      235                      240                      245                      250

tcg ttg cca ggt gaa gct tat tgg gga acg caa cga ccg tct ttc taa 879
Ser Leu Pro Gly Glu Ala Tyr Trp Gly Thr Gln Arg Pro Ser Phe *
      255                      260                      265

gataatatca ttgataatat aagtttcgtc ttcttattct ttttcacttt ttaccttttt 939
cactttctta ggttttgttt taacgtttga ttaatacctg aagggttttg gaaaattttc 999
gatcgataaa aaggatttat gttgcgagcc gaaacgcggc c 1040

```

<210> 92  
 <211> 265  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (53)...(73)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (121)...(165)  
 <223> Conserved domain

<400> 92  
 Met Asp Glu Val Ser Arg Ser His Thr Pro Gln Phe Leu Ser Ser Asp  
 1 5 10 15  
 His Gln His Tyr His His Gln Asn Ala Gly Arg Gln Lys Arg Gly Arg  
 20 25 30  
 Glu Glu Glu Gly Val Glu Pro Asn Ile Gly Glu Asp Leu Ala Thr  
 35 40 45  
 Phe Pro Ser Gly Glu Glu Asn Ile Lys Lys Arg Arg Pro Arg Gly Arg  
 50 55 60  
 Pro Ala Gly Ser Lys Asn Lys Pro Lys Ala Pro Ile Ile Val Thr Arg  
 65 70 75 80  
 Asp Ser Ala Asn Ala Phe Arg Cys His Val Met Glu Ile Thr Asn Ala  
 85 90 95  
 Cys Asp Val Met Glu Ser Leu Ala Val Phe Ala Arg Arg Arg Gln Arg

```

      100      105      110
Gly Val Cys Val Leu Thr Gly Asn Gly Ala Val Thr Asn Val Thr Val
      115      120      125
Arg Gln Pro Gly Gly Gly Val Val Ser Leu His Gly Arg Phe Glu Ile
      130      135      140
Leu Ser Leu Ser Gly Ser Phe Leu Pro Pro Pro Ala Pro Pro Ala Ala
145      150      155      160
Ser Gly Leu Lys Val Tyr Leu Ala Gly Gly Gln Gly Gln Val Ile Gly
      165      170      175
Gly Ser Val Val Gly Pro Leu Thr Ala Ser Ser Pro Val Val Val Met
      180      185      190
Ala Ala Ser Phe Gly Asn Ala Ser Tyr Glu Arg Leu Pro Leu Glu Glu
      195      200      205
Glu Glu Glu Thr Glu Arg Glu Ile Asp Gly Asn Ala Ala Arg Ala Ile
      210      215      220
Gly Thr Gln Thr Gln Lys Gln Leu Met Gln Asp Ala Thr Ser Phe Ile
225      230      235      240
Gly Ser Pro Ser Asn Leu Ile Asn Ser Val Ser Leu Pro Gly Glu Ala
      245      250      255
Tyr Trp Gly Thr Gln Arg Pro Ser Phe
      260      265

```

&lt;210&gt; 93

&lt;211&gt; 1072

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (87)...(977)

&lt;400&gt; 93

```

aaaagataac gtagccatta acttttatgc attaactcct tcatttcttt ttgtgctcgt 60
ttggttgaga gagagagagg ctatat atg atg ttc caa caa gat tac cct cat 113
                        Met Met Phe Gln Gln Asp Tyr Pro His
                        1                      5

```

```

ggc ttt tca ctc gtg gaa aca tcc tta agt tac gaa atg ttg gat tac 161
Gly Phe Ser Leu Val Glu Thr Ser Leu Ser Tyr Glu Met Leu Asp Tyr
10                      15                      20                      25

```

```

ttt caa aac atc gtc gtt tcg aac tct gaa gac gtg gcg tca cag caa 209
Phe Gln Asn Ile Val Val Ser Asn Ser Glu Asp Val Ala Ser Gln Gln
30                      35                      40

```

```

aat tcc att tcg tcc tct tct tat tca tca gcg aca ctc tcc tgc tcc 257
Asn Ser Ile Ser Ser Ser Ser Tyr Ser Ser Ala Thr Leu Ser Cys Ser
45                      50                      55

```

```

ata aca gag caa aaa tct cac tta act gaa aag tta tct cct cta cga 305
Ile Thr Glu Gln Lys Ser His Leu Thr Glu Lys Leu Ser Pro Leu Arg
60                      65                      70

```

```

gaa aga tat ggt tgc ggt gac ttt ctg tcg cgg aag agg aga agg aga 353
Glu Arg Tyr Gly Cys Gly Asp Phe Leu Ser Arg Lys Arg Arg Arg Arg
75                      80                      85

```

```

agt gaa aaa acg att gta gat aaa gag aat caa agg atg aat cac att 401
Ser Glu Lys Thr Ile Val Asp Lys Glu Asn Gln Arg Met Asn His Ile
90 95 100 105

gcc gtc gag cgt aac cgg aga aaa cag atg aat cat ttt ctg tct atc 449
Ala Val Glu Arg Asn Arg Arg Lys Gln Met Asn His Phe Leu Ser Ile
110 115 120

ctc aag tct atg atg cct ctc tct tat tct caa cct aat gac caa gca 497
Leu Lys Ser Met Met Pro Leu Ser Tyr Ser Gln Pro Asn Asp Gln Ala
125 130 135

tca atc ata gaa ggg acc att agc tat ctg aag aag cta gaa caa cgt 545
Ser Ile Ile Glu Gly Thr Ile Ser Tyr Leu Lys Lys Leu Glu Gln Arg
140 145 150

ctc caa tct ctc gaa gcc caa tta aaa gct act aaa ctc aat caa tca 593
Leu Gln Ser Leu Glu Ala Gln Leu Lys Ala Thr Lys Leu Asn Gln Ser
155 160 165

cca aat ata ttt tcc gac ttc ttc atg ttc cct caa tac tcc acc gcc 641
Pro Asn Ile Phe Ser Asp Phe Phe Met Phe Pro Gln Tyr Ser Thr Ala
170 175 180 185

act gcc act gcc acc gcc act gcc tcc tca tcc tcc tcg agc cac cac 689
Thr Ala Thr Ala Thr Ala Thr Ala Ser Ser Ser Ser Ser Ser His His
190 195 200

cat cac aag cga cta gag gtg gtt gct gac gtg gag gtt aca atg gta 737
His His Lys Arg Leu Glu Val Val Ala Asp Val Glu Val Thr Met Val
205 210 215

gaa aga cat gcc aac att aaa gtg tta acg aag aca cag cca aga ttg 785
Glu Arg His Ala Asn Ile Lys Val Leu Thr Lys Thr Gln Pro Arg Leu
220 225 230

ctc ttc aag att atc aat gag ttt aac tct tta ggt tta agt act ctt 833
Leu Phe Lys Ile Ile Asn Glu Phe Asn Ser Leu Gly Leu Ser Thr Leu
235 240 245

cat ctc aac ctc aca act tcc aaa gac atg tct ctc ttc act ttt agc 881
His Leu Asn Leu Thr Thr Ser Lys Asp Met Ser Leu Phe Thr Phe Ser
250 255 260 265

gtc aag gta gag gca gat tgt caa ttg acg cct tct ggt aat gag gtc 929
Val Lys Val Glu Ala Asp Cys Gln Leu Thr Pro Ser Gly Asn Glu Val
270 275 280

gca aat acg gtg cat gaa gtc gtt aga aga gtt cac aag gaa cgt tga 977
Ala Asn Thr Val His Glu Val Val Arg Arg Val His Lys Glu Arg *
285 290 295

atctttgttta catactagct aactttgaaa ttctatttta ttgtataaac aatctcttat 1037
gtgtgtaatt tacatatata caattaatta acatt 1072

```

&lt;210&gt; 94

<211> 296  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (92)...(161)  
 <223> Conserved domain

<400> 94  
 Met Met Phe Gln Gln Asp Tyr Pro His Gly Phe Ser Leu Val Glu Thr  
 1 5 10 15  
 Ser Leu Ser Tyr Glu Met Leu Asp Tyr Phe Gln Asn Ile Val Val Ser  
 20 25 30  
 Asn Ser Glu Asp Val Ala Ser Gln Gln Asn Ser Ile Ser Ser Ser Ser  
 35 40 45  
 Tyr Ser Ser Ala Thr Leu Ser Cys Ser Ile Thr Glu Gln Lys Ser His  
 50 55 60  
 Leu Thr Glu Lys Leu Ser Pro Leu Arg Glu Arg Tyr Gly Cys Gly Asp  
 65 70 75 80  
 Phe Leu Ser Arg Lys Arg Arg Arg Arg Ser Glu Lys Thr Ile Val Asp  
 85 90 95  
 Lys Glu Asn Gln Arg Met Asn His Ile Ala Val Glu Arg Asn Arg Arg  
 100 105 110  
 Lys Gln Met Asn His Phe Leu Ser Ile Leu Lys Ser Met Met Pro Leu  
 115 120 125  
 Ser Tyr Ser Gln Pro Asn Asp Gln Ala Ser Ile Ile Glu Gly Thr Ile  
 130 135 140  
 Ser Tyr Leu Lys Lys Leu Glu Gln Arg Leu Gln Ser Leu Glu Ala Gln  
 145 150 155 160  
 Leu Lys Ala Thr Lys Leu Asn Gln Ser Pro Asn Ile Phe Ser Asp Phe  
 165 170 175  
 Phe Met Phe Pro Gln Tyr Ser Thr Ala Thr Ala Thr Ala Thr  
 180 185 190  
 Ala Ser Ser Ser Ser Ser His His His Lys Arg Leu Glu Val  
 195 200 205  
 Val Ala Asp Val Glu Val Thr Met Val Glu Arg His Ala Asn Ile Lys  
 210 215 220  
 Val Leu Thr Lys Thr Gln Pro Arg Leu Leu Phe Lys Ile Ile Asn Glu  
 225 230 235 240  
 Phe Asn Ser Leu Gly Leu Ser Thr Leu His Leu Asn Leu Thr Thr Ser  
 245 250 255  
 Lys Asp Met Ser Leu Phe Thr Phe Ser Val Lys Val Glu Ala Asp Cys  
 260 265 270  
 Gln Leu Thr Pro Ser Gly Asn Glu Val Ala Asn Thr Val His Glu Val  
 275 280 285  
 Val Arg Arg Val His Lys Glu Arg  
 290 295

<210> 95  
 <211> 876  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS

&lt;222&gt; (1)...(876)

&lt;400&gt; 95

atg gag gct cga tgt gac ttt tgt gga acc gaa aaa gca tta atc tac	48
Met Glu Ala Arg Cys Asp Phe Cys Gly Thr Glu Lys Ala Leu Ile Tyr	
1 5 10 15	
tgc aaa tct gat tgc gct aaa ctt tgt tta aac tgt gac gtt aat gta	96
Cys Lys Ser Asp Ser Ala Lys Leu Cys Leu Asn Cys Asp Val Asn Val	
20 25 30	
cac tca gcg aat cct ctg tct caa aga cac acg cga tct cta ctc tgc	144
His Ser Ala Asn Pro Leu Ser Gln Arg His Thr Arg Ser Leu Leu Cys	
35 40 45	
gag aaa tgt tcc tta caa ccc acg gca gta cac tgt atg aac gaa aat	192
Glu Lys Cys Ser Leu Gln Pro Thr Ala Val His Cys Met Asn Glu Asn	
50 55 60	
gtt tct ctt tgc caa gga tgt cag tgg acc gct agt aac tgc act ggt	240
Val Ser Leu Cys Gln Gly Cys Gln Trp Thr Ala Ser Asn Cys Thr Gly	
65 70 75 80	
tta gga cac aga ctt cag agt ctg aac cct tac tct gat tgt cct tct	288
Leu Gly His Arg Leu Gln Ser Leu Asn Pro Tyr Ser Asp Cys Pro Ser	
85 90 95	
cca tca gat ttc ggc aag atc tgg tct tca act tta gaa cct tct gtc	336
Pro Ser Asp Phe Gly Lys Ile Trp Ser Ser Thr Leu Glu Pro Ser Val	
100 105 110	
aca agt ttg gtc tct cca ttt tct gat acc ctg ctc cag gaa tta gac	384
Thr Ser Leu Val Ser Pro Phe Ser Asp Thr Leu Leu Gln Glu Leu Asp	
115 120 125	
gat tgg aat ggt tca tca act tcc gtg gtt aca caa act cag aac ctt	432
Asp Trp Asn Gly Ser Ser Thr Ser Val Val Thr Gln Thr Gln Asn Leu	
130 135 140	
aaa gac tat tct tct ttc ttt cct atg gaa tct aat ctt cca aaa gtg	480
Lys Asp Tyr Ser Ser Phe Phe Pro Met Glu Ser Asn Leu Pro Lys Val	
145 150 155 160	
ata gaa gag gaa tgt tct ggt ttg gat cta tgc gaa ggg ata aac tta	528
Ile Glu Glu Glu Cys Ser Gly Leu Asp Leu Cys Glu Gly Ile Asn Leu	
165 170 175	
gat gat gcg ccg ttg aat ttc aac gca agc aat gat ata atc gga tgt	576
Asp Asp Ala Pro Leu Asn Phe Asn Ala Ser Asn Asp Ile Ile Gly Cys	
180 185 190	
tca tca ctt gac aat acc aaa tgt tac gag tac gag gat tct ttt aag	624
Ser Ser Leu Asp Asn Thr Lys Cys Tyr Glu Tyr Glu Asp Ser Phe Lys	
195 200 205	
gaa gag aat aac att ggc ctt cct tca tta ctg ctt cct act tta tcc	672
Glu Glu Asn Asn Ile Gly Leu Pro Ser Leu Leu Leu Pro Thr Leu Ser	

```

210                215                220
ggg aac gtt gta ccg aac atg tca ctt tcg atg tct aac ctc aca gga 720
Gly Asn Val Val Pro Asn Met Ser Leu Ser Met Ser Asn Leu Thr Gly
225                230                235                240

gaa agc aat gcg act gat tac caa gat tgc gga att tca cca ggg ttt 768
Glu Ser Asn Ala Thr Asp Tyr Gln Asp Cys Gly Ile Ser Pro Gly Phe
245                250                255

ctc ata gga gat tcg cca tgg gaa tcg aat gtt gaa gtt agt ttt aat 816
Leu Ile Gly Asp Ser Pro Trp Glu Ser Asn Val Glu Val Ser Phe Asn
260                265                270

cca aaa tta agg gat gag gct aaa aag aga tac aaa caa aag aaa tca 864
Pro Lys Leu Arg Asp Glu Ala Lys Lys Arg Tyr Lys Gln Lys Lys Ser
275                280                285

aag cgc atg taa 876
Lys Arg Met *
290

```

<210> 96  
 <211> 291  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (5)...(28)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (48)...(71)  
 <223> Conserved domain

```

<400> 96
Met Glu Ala Arg Cys Asp Phe Cys Gly Thr Glu Lys Ala Leu Ile Tyr
 1          5          10         15
Cys Lys Ser Asp Ser Ala Lys Leu Cys Leu Asn Cys Asp Val Asn Val
20         25         30
His Ser Ala Asn Pro Leu Ser Gln Arg His Thr Arg Ser Leu Leu Cys
35         40         45
Glu Lys Cys Ser Leu Gln Pro Thr Ala Val His Cys Met Asn Glu Asn
50         55         60
Val Ser Leu Cys Gln Gly Cys Gln Trp Thr Ala Ser Asn Cys Thr Gly
65         70         75         80
Leu Gly His Arg Leu Gln Ser Leu Asn Pro Tyr Ser Asp Cys Pro Ser
85         90         95
Pro Ser Asp Phe Gly Lys Ile Trp Ser Ser Thr Leu Glu Pro Ser Val
100        105        110
Thr Ser Leu Val Ser Pro Phe Ser Asp Thr Leu Leu Gln Glu Leu Asp
115        120        125
Asp Trp Asn Gly Ser Ser Thr Ser Val Val Thr Gln Thr Gln Asn Leu
130        135        140
Lys Asp Tyr Ser Ser Phe Phe Pro Met Glu Ser Asn Leu Pro Lys Val

```

145                      150                      155                      160  
 Ile Glu Glu Glu Cys Ser Gly Leu Asp Leu Cys Glu Gly Ile Asn Leu  
                                  165                      170                      175  
 Asp Asp Ala Pro Leu Asn Phe Asn Ala Ser Asn Asp Ile Ile Gly Cys  
                                  180                      185                      190  
 Ser Ser Leu Asp Asn Thr Lys Cys Tyr Glu Tyr Glu Asp Ser Phe Lys  
                                  195                      200                      205  
 Glu Glu Asn Asn Ile Gly Leu Pro Ser Leu Leu Leu Pro Thr Leu Ser  
                                  210                      215                      220  
 Gly Asn Val Val Pro Asn Met Ser Leu Ser Met Ser Asn Leu Thr Gly  
 225                                   230                                   235                                   240  
 Glu Ser Asn Ala Thr Asp Tyr Gln Asp Cys Gly Ile Ser Pro Gly Phe  
                                  245                                   250                                   255  
 Leu Ile Gly Asp Ser Pro Trp Glu Ser Asn Val Glu Val Ser Phe Asn  
                                  260                                   265                                   270  
 Pro Lys Leu Arg Asp Glu Ala Lys Lys Arg Tyr Lys Gln Lys Lys Ser  
                                  275                                   280                                   285  
 Lys Arg Met  
                                  290

<210> 97  
 <211> 727  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (82)...(570)

<400> 97  
 accaaactca aaaaacacaa accacaagag gatcatttca ttttttattg tttcgtttta 60  
 atcatcatca tcagaagaaa a atg gtt gcg ata tcg gag atc aag tcg acg 111  
                                  Met Val Ala Ile Ser Glu Ile Lys Ser Thr  
                                  1                                   5                                   10  
  
 gtg gat gtc acg gcg gcg aat tgt ttg atg ctt tta tct aga gtt gga 159  
 Val Asp Val Thr Ala Ala Asn Cys Leu Met Leu Leu Ser Arg Val Gly  
                                  15                                   20                                   25  
  
 caa gaa aac gtt gac ggt ggc gat caa aaa cgc gtt ttc aca tgt aaa 207  
 Gln Glu Asn Val Asp Gly Gly Asp Gln Lys Arg Val Phe Thr Cys Lys  
                                  30                                   35                                   40  
  
 acg tgt ttg aag cag ttt cat tcg ttc caa gcc tta gga ggt cac cgt 255  
 Thr Cys Leu Lys Gln Phe His Ser Phe Gln Ala Leu Gly Gly His Arg  
                                  45                                   50                                   55  
  
 gcg agt cac aag aag cct aac aac gac gct ttg tcg tct gga ttg atg 303  
 Ala Ser His Lys Lys Pro Asn Asn Asp Ala Leu Ser Ser Gly Leu Met  
                                  60                                   65                                   70  
  
 aag aag gtg aaa acg tcg tcg cat cct tgt ccc ata tgt gga gtg gag 351  
 Lys Lys Val Lys Thr Ser Ser His Pro Cys Pro Ile Cys Gly Val Glu  
                                  75                                   80                                   85                                   90  
  
 ttt ccg atg gga caa gct ttg gga gga cac atg agg aga cac agg aac 399

```

Phe Pro Met Gly Gln Ala Leu Gly Gly His Met Arg Arg His Arg Asn
          95                      100                      105

gag agt ggg gct gct ggt ggc gcg ttg gtt aca cgc gct ttg ttg ccg   447
Glu Ser Gly Ala Ala Gly Gly Ala Leu Val Thr Arg Ala Leu Leu Pro
          110                      115                      120

gag ccc acg gtg act acg ttg aag aaa tct agc agt ggg aag aga gtg   495
Glu Pro Thr Val Thr Thr Leu Lys Lys Ser Ser Ser Gly Lys Arg Val
          125                      130                      135

gct tgt ttg gat ctg agt cta ggg atg gtg gac aat ttg aat ctc aag   543
Ala Cys Leu Asp Leu Ser Leu Gly Met Val Asp Asn Leu Asn Leu Lys
          140                      145                      150

ttg gag ctt gga aga aca gtt tat tga ttttatttat tttccttaaa       590
Leu Glu Leu Gly Arg Thr Val Tyr *
          155                      160

ttttctgaat atatttggtt ctctcattct ttgaattttt cttaatatcc tagattatac 650
atacatccgc agatttagga aactttcata gagtgtaatc ttttctttct gtaaaaatat 710
attttacttg tagcaaa                                           727

```

```

<210> 98
<211> 162
<212> PRT
<213> Arabidopsis thaliana

```

```

<220>
<221> DOMAIN
<222> (41)...(61)
<223> Conserved domain

```

```

<221> DOMAIN
<222> (84)...(104)
<223> Conserved domain

```

```

<400> 98
Met Val Ala Ile Ser Glu Ile Lys Ser Thr Val Asp Val Thr Ala Ala
 1          5          10          15
Asn Cys Leu Met Leu Leu Ser Arg Val Gly Gln Glu Asn Val Asp Gly
          20          25          30
Gly Asp Gln Lys Arg Val Phe Thr Cys Lys Thr Cys Leu Lys Gln Phe
          35          40          45
His Ser Phe Gln Ala Leu Gly Gly His Arg Ala Ser His Lys Lys Pro
          50          55          60
Asn Asn Asp Ala Leu Ser Ser Gly Leu Met Lys Lys Val Lys Thr Ser
          65          70          75          80
Ser His Pro Cys Pro Ile Cys Gly Val Glu Phe Pro Met Gly Gln Ala
          85          90          95
Leu Gly Gly His Met Arg Arg His Arg Asn Glu Ser Gly Ala Ala Gly
          100          105          110
Gly Ala Leu Val Thr Arg Ala Leu Leu Pro Glu Pro Thr Val Thr Thr
          115          120          125
Leu Lys Lys Ser Ser Ser Gly Lys Arg Val Ala Cys Leu Asp Leu Ser
          130          135          140
Leu Gly Met Val Asp Asn Leu Asn Leu Lys Leu Glu Leu Gly Arg Thr

```



```

145                               150                               155                               160
Val Tyr

<210> 99
<211> 785
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (38)...(472)

<400> 99
cttcttcttc acatcgatca tcatacaaca acaaaaa atg gat tac aga gaa tcc 55
                               Met Asp Tyr Arg Glu Ser
                               1                               5

acc ggt gaa agt cag tca aag tac aaa gga atc cgt cgt cgg aaa tgg 103
Thr Gly Glu Ser Gln Ser Lys Tyr Lys Gly Ile Arg Arg Arg Lys Trp
          10                               15                               20

ggc aaa tgg gta tca gag att aga gtt ccg gga act cgt gac cgt ctc 151
Gly Lys Trp Val Ser Glu Ile Arg Val Pro Gly Thr Arg Asp Arg Leu
          25                               30                               35

tgg tta ggt tca ttc tca aca gca gaa ggt gcc gcc gta gca cac gac 199
Trp Leu Gly Ser Phe Ser Thr Ala Glu Gly Ala Ala Val Ala His Asp
          40                               45                               50

gtt gct ttc ttc tgt tta cac caa cct gat tct tta gaa tct ctc aat 247
Val Ala Phe Phe Cys Leu His Gln Pro Asp Ser Leu Glu Ser Leu Asn
          55                               60                               65                               70

ttc cct cat ttg ctt aat cct tca ctc gtt tcc aga act tct ccg aga 295
Phe Pro His Leu Leu Asn Pro Ser Leu Val Ser Arg Thr Ser Pro Arg
          75                               80                               85

tct atc cag caa gct gct tct aac gcc ggc atg gcc att gac gcc gga 343
Ser Ile Gln Gln Ala Ala Ser Asn Ala Gly Met Ala Ile Asp Ala Gly
          90                               95                               100

atc gtc cac agt acc agc gtg aac tct gga tgc gga gat acg acg acg 391
Ile Val His Ser Thr Ser Val Asn Ser Gly Cys Gly Asp Thr Thr Thr
          105                               110                               115

tat tac gag aat gga gct gat caa gtg gag ccg ttg aat att tca gtg 439
Tyr Tyr Glu Asn Gly Ala Asp Gln Val Glu Pro Leu Asn Ile Ser Val
          120                               125                               130

tat gat tat ctg ggc ggc cac gat cac gtt tga tttatctcga cggatcatgat 492
Tyr Asp Tyr Leu Gly Gly His Asp His Val *
135                               140

cacgtttgat cttcttttga gtaagat ttt gtaccataat caaacacaggt gtggtgctaa 552
aatcttactc aaaacaagat taggtaccac agagaaacaa tcaaatggtt gtgaatatac 612

```

attataaggt ttgattaat gtttgtttca ctgatttagt gaagtttggt ccattgtata 672  
 caaatctatt caagaaacct agcgcgagat catgtttcgt gattgaagat tgagattttt 732  
 aagtattcgt aatatttttg taaaatacaa ataaaaaaaa aaaaaaaaaa aaa 785

<210> 100  
 <211> 144  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (11)...(80)  
 <223> Conserved domain

<400> 100  
 Met Asp Tyr Arg Glu Ser Thr Gly Glu Ser Gln Ser Lys Tyr Lys Gly  
 1 5 10 15  
 Ile Arg Arg Arg Lys Trp Gly Lys Trp Val Ser Glu Ile Arg Val Pro  
 20 25 30  
 Gly Thr Arg Asp Arg Leu Trp Leu Gly Ser Phe Ser Thr Ala Glu Gly  
 35 40 45  
 Ala Ala Val Ala His Asp Val Ala Phe Phe Cys Leu His Gln Pro Asp  
 50 55 60  
 Ser Leu Glu Ser Leu Asn Phe Pro His Leu Leu Asn Pro Ser Leu Val  
 65 70 75 80  
 Ser Arg Thr Ser Pro Arg Ser Ile Gln Gln Ala Ala Ser Asn Ala Gly  
 85 90 95  
 Met Ala Ile Asp Ala Gly Ile Val His Ser Thr Ser Val Asn Ser Gly  
 100 105 110  
 Cys Gly Asp Thr Thr Thr Tyr Tyr Glu Asn Gly Ala Asp Gln Val Glu  
 115 120 125  
 Pro Leu Asn Ile Ser Val Tyr Asp Tyr Leu Gly Gly His Asp His Val  
 130 135 140

<210> 101  
 <211> 1008  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (33)...(695)

<400> 101  
 tggatcaaca agaccatgga cagtctggag ct atg aac tat ggc aca aac cca 53  
 Met Asn Tyr Gly Thr Asn Pro  
 1 5  
 tac caa acc aac ccg atg agc acc act gct gct act gta gca gga ggt 101  
 Tyr Gln Thr Asn Pro Met Ser Thr Thr Ala Ala Thr Val Ala Gly Gly  
 10 15 20  
 gcg gca caa cca ggc cag ctg gcg ttc cac cag atc cat cag cag cag 149  
 Ala Ala Gln Pro Gly Gln Leu Ala Phe His Gln Ile His Gln Gln Gln  
 25 30 35

```

cag cag caa cag ctg gca cag cag ctt caa gca ttt tgg gag aac caa 197
Gln Gln Gln Gln Leu Ala Gln Gln Leu Gln Ala Phe Trp Glu Asn Gln
40 45 50 55

ttc aaa gag att gag aag act acc gat ttc aag aac cac agc ctt ccc 245
Phe Lys Glu Ile Glu Lys Thr Thr Asp Phe Lys Asn His Ser Leu Pro
60 65 70

ctt gcg aga atc aag aaa atc atg aaa gcg gat gaa gat gtc cgt atg 293
Leu Ala Arg Ile Lys Lys Ile Met Lys Ala Asp Glu Asp Val Arg Met
75 80 85

atc tcg gct gag gcg ccg gtc gtg ttt gca agg gcc tgt gag atg ttc 341
Ile Ser Ala Glu Ala Pro Val Val Phe Ala Arg Ala Cys Glu Met Phe
90 95 100

atc ctg gag ctg aca ctc agg tcg tgg aac cac act gag gag aat aag 389
Ile Leu Glu Leu Thr Leu Arg Ser Trp Asn His Thr Glu Glu Asn Lys
105 110 115

agg cgg acg ttg cag aag aac gat att gct gct gct gtg act aga acc 437
Arg Arg Thr Leu Gln Lys Asn Asp Ile Ala Ala Val Thr Arg Thr
120 125 130 135

gat att ttt gat ttc ctt gtg gat att gtt ccc cgg gag gat ctc cga 485
Asp Ile Phe Asp Phe Leu Val Asp Ile Val Pro Arg Glu Asp Leu Arg
140 145 150

gat gaa gtc ttg gga agt att ccg agg ggc act gtc ccg gaa gct gct 533
Asp Glu Val Leu Gly Ser Ile Pro Arg Gly Thr Val Pro Glu Ala Ala
155 160 165

gct gct ggt tac ccg tat gga tac ttg cct gca gga act gct cca ata 581
Ala Ala Gly Tyr Pro Tyr Gly Tyr Leu Pro Ala Gly Thr Ala Pro Ile
170 175 180

gga aat ccg gga atg gtt atg ggt aat ccc ggt ggt gcg tat cca cct 629
Gly Asn Pro Gly Met Val Met Gly Asn Pro Gly Gly Ala Tyr Pro Pro
185 190 195

aat cct tat atg ggt caa cca atg tgg caa caa cag gca cct gac caa 677
Asn Pro Tyr Met Gly Gln Pro Met Trp Gln Gln Ala Pro Asp Gln
200 205 210 215

cct gac cag gaa aat tag caagaaactg tgagtcttcc agcttcgcgg 725
Pro Asp Gln Glu Asn *
220

ccgctctaga caggcctcgt accggatcct ctagctagag ctttcggtcg tatcatcggt 785
ttcgacaacg ttcgtcaagt tcaatgcac agtttcattg cgcacacacc agaactctac 845
tgagtttgag tattatggca ttgggaaaac tgtttttctt gtccatttgt tgtgcttgta 905
atttactgtg ttttttattc gggtttcgtc atcgaactgt gaaatggaaa tggatggaga 965
agagttaatg aatgatatgg ccttttgctc attctcaaat taa 1008

<210> 102
<211> 220
<212> PRT

```

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (57)...(156)

<223> Conserved domain

<400> 102

```

Met Asn Tyr Gly Thr Asn Pro Tyr Gln Thr Asn Pro Met Ser Thr Thr
 1           5           10           15
Ala Ala Thr Val Ala Gly Gly Ala Ala Gln Pro Gly Gln Leu Ala Phe
          20           25           30
His Gln Ile His Gln Gln Gln Gln Gln Gln Leu Ala Gln Gln Leu
          35           40           45
Gln Ala Phe Trp Glu Asn Gln Phe Lys Glu Ile Glu Lys Thr Thr Asp
          50           55           60
Phe Lys Asn His Ser Leu Pro Leu Ala Arg Ile Lys Lys Ile Met Lys
          65           70           75           80
Ala Asp Glu Asp Val Arg Met Ile Ser Ala Glu Ala Pro Val Val Phe
          85           90           95
Ala Arg Ala Cys Glu Met Phe Ile Leu Glu Leu Thr Leu Arg Ser Trp
          100          105          110
Asn His Thr Glu Glu Asn Lys Arg Arg Thr Leu Gln Lys Asn Asp Ile
          115          120          125
Ala Ala Ala Val Thr Arg Thr Asp Ile Phe Asp Phe Leu Val Asp Ile
          130          135          140
Val Pro Arg Glu Asp Leu Arg Asp Glu Val Leu Gly Ser Ile Pro Arg
          145          150          155          160
Gly Thr Val Pro Glu Ala Ala Ala Ala Gly Tyr Pro Tyr Gly Tyr Leu
          165          170          175
Pro Ala Gly Thr Ala Pro Ile Gly Asn Pro Gly Met Val Met Gly Asn
          180          185          190
Pro Gly Gly Ala Tyr Pro Pro Asn Pro Tyr Met Gly Gln Pro Met Trp
          195          200          205
Gln Gln Gln Ala Pro Asp Gln Pro Asp Gln Glu Asn
          210          215          220

```

<210> 103

<211> 1409

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (224)...(1093)

<400> 103

```

ttgatgccgc tcaatccac tctcttcgc aaggaccctt cctctatata aggaagttca 60
tttcatttgg agaggacacg ctgacaagct gactctagca gatctgggac cgtcgacca 120
cgcgctccgaa ttgattagga taggatcagg atcatcctca acaacctcct cctaattcct 180
cctccattca tagtaacaat aatattaaga aagagggtaa act atg tca gaa tta 235
                                     Met Ser Glu Leu
                                     1

```

```

tta cag ttg cct cca ggt ttc cga ttt cac cct acc gat gaa gag ctt 283
Leu Gln Leu Pro Pro Gly Phe Arg Phe His Pro Thr Asp Glu Glu Leu

```

5	10	15	20	
gtc atg cac tat ctc tgc cgc aaa tgt gcc tct cag tcc atc gcc gtt	331			
Val Met His Tyr Leu Cys Arg Lys Cys Ala Ser Gln Ser Ile Ala Val				
25 30 35				
ccg atc atc gct gag atc gat ctc tac aaa tac gat cca tgg gag ctt	379			
Pro Ile Ile Ala Glu Ile Asp Leu Tyr Lys Tyr Asp Pro Trp Glu Leu				
40 45 50				
cct ggt tta gcc ttg tat ggt gag aag gaa tgg tac ttc ttc tct ccc	427			
Pro Gly Leu Ala Leu Tyr Gly Glu Lys Glu Trp Tyr Phe Phe Ser Pro				
55 60 65				
agg gac aga aaa tat ccc aac ggt tgc cgt cct aac cgg tcc gct ggt	475			
Arg Asp Arg Lys Tyr Pro Asn Gly Ser Arg Pro Asn Arg Ser Ala Gly				
70 75 80				
tct ggt tac tgg aaa gct acc gga gct gat aaa ccg atc gga cta cct	523			
Ser Gly Tyr Trp Lys Ala Thr Gly Ala Asp Lys Pro Ile Gly Leu Pro				
85 90 95 100				
aaa ccg gtc gga att aag aaa gct ctt gtt ttc tac gcc ggc aaa gct	571			
Lys Pro Val Gly Ile Lys Lys Ala Leu Val Phe Tyr Ala Gly Lys Ala				
105 110 115				
cca aag gga gag aaa acc aat tgg atc atg cac gag tac cgt ctc gcc	619			
Pro Lys Gly Glu Lys Thr Asn Trp Ile Met His Glu Tyr Arg Leu Ala				
120 125 130				
gac gtt gac cgg tcc gtt cgc aag aag aag aat agt ctc agg ctg gat	667			
Asp Val Asp Arg Ser Val Arg Lys Lys Lys Asn Ser Leu Arg Leu Asp				
135 140 145				
gat tgg gtt ctc tgc cgg att tac aac aaa aaa gga gct acc gag agg	715			
Asp Trp Val Leu Cys Arg Ile Tyr Asn Lys Lys Gly Ala Thr Glu Arg				
150 155 160				
cgg gga cca ccg cct ccg gtt gtt tac ggc gac gaa atc atg gag gag	763			
Arg Gly Pro Pro Pro Val Val Tyr Gly Asp Glu Ile Met Glu Glu				
165 170 175 180				
aag ccg aag gtg acg gag atg gtt atg cct ccg ccg ccg caa cag aca	811			
Lys Pro Lys Val Thr Glu Met Val Met Pro Pro Pro Pro Gln Gln Thr				
185 190 195				
agt gag ttc gcg tat ttc gac acg tgc gat tgc gtg ccg aag ctg cat	859			
Ser Glu Phe Ala Tyr Phe Asp Thr Ser Asp Ser Val Pro Lys Leu His				
200 205 210				
act acg gat tgc agt tgc tgc gag cag gtg gtg tgc ccg gag ttc acg	907			
Thr Thr Asp Ser Ser Cys Ser Glu Gln Val Val Ser Pro Glu Phe Thr				
215 220 225				
agc gag gtt cag agc gag ccc aag tgg aaa gat tgg tgc gcc gta agt	955			
Ser Glu Val Gln Ser Glu Pro Lys Trp Lys Asp Trp Ser Ala Val Ser				
230 235 240				

```

aat gac aat aac aat acc ctt gat ttt ggg ttt aat tac att gat gcc 1003
Asn Asp Asn Asn Asn Thr Leu Asp Phe Gly Phe Asn Tyr Ile Asp Ala
245                250                255                260

acc gtg gat aac gcg ttt gga gga gga ggg agt agt aat cag atg ttt 1051
Thr Val Asp Asn Ala Phe Gly Gly Gly Gly Ser Ser Asn Gln Met Phe
                265                270                275

ccg cta cag gat atg ttc atg tac atg cag aag cct tac tag 1093
Pro Leu Gln Asp Met Phe Met Tyr Met Gln Lys Pro Tyr *
                280                285

aaggaattc ctttcctgcc gccgaaacgc aacgcaaac gacctcgtt tttgcgttta 1153
tggcaacacg agaccgtttt atatggtcaa tgagtgtgcc gattcggcca ttagatttct 1213
gttcagtctt cgtttattct atagaccgtc cgatttcaga tcatccctaa tcggacggtg 1273
gtcgttggat gtatcagtag tgtattactg tgttaggtag aagaaaatcc acttggttctt 1333
aaattggcat aaaagtcaga agctaataatt tatatgtgcc gcaatcaatt taatattttc 1393
tgtctaaaaa aaaaaa 1409

```

<210> 104  
 <211> 289  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (10)...(155)  
 <223> Conserved domain

<400> 104

```

Met Ser Glu Leu Leu Gln Leu Pro Pro Gly Phe Arg Phe His Pro Thr
1          5          10          15
Asp Glu Glu Leu Val Met His Tyr Leu Cys Arg Lys Cys Ala Ser Gln
20        25        30
Ser Ile Ala Val Pro Ile Ile Ala Glu Ile Asp Leu Tyr Lys Tyr Asp
35        40        45
Pro Trp Glu Leu Pro Gly Leu Ala Leu Tyr Gly Glu Lys Glu Trp Tyr
50        55        60
Phe Phe Ser Pro Arg Asp Arg Lys Tyr Pro Asn Gly Ser Arg Pro Asn
65        70        75        80
Arg Ser Ala Gly Ser Gly Tyr Trp Lys Ala Thr Gly Ala Asp Lys Pro
85        90        95
Ile Gly Leu Pro Lys Pro Val Gly Ile Lys Lys Ala Leu Val Phe Tyr
100       105       110
Ala Gly Lys Ala Pro Lys Gly Glu Lys Thr Asn Trp Ile Met His Glu
115       120       125
Tyr Arg Leu Ala Asp Val Asp Arg Ser Val Arg Lys Lys Lys Asn Ser
130       135       140
Leu Arg Leu Asp Asp Trp Val Leu Cys Arg Ile Tyr Asn Lys Lys Gly
145       150       155       160
Ala Thr Glu Arg Arg Gly Pro Pro Pro Pro Val Val Tyr Gly Asp Glu
165       170       175
Ile Met Glu Glu Lys Pro Lys Val Thr Glu Met Val Met Pro Pro Pro
180       185       190
Pro Gln Gln Thr Ser Glu Phe Ala Tyr Phe Asp Thr Ser Asp Ser Val
195       200       205

```

Pro Lys Leu His Thr Thr Asp Ser Ser Cys Ser Glu Gln Val Val Ser  
 210 215 220  
 Pro Glu Phe Thr Ser Glu Val Gln Ser Glu Pro Lys Trp Lys Asp Trp  
 225 230 235 240  
 Ser Ala Val Ser Asn Asp Asn Asn Asn Thr Leu Asp Phe Gly Phe Asn  
 245 250 255  
 Tyr Ile Asp Ala Thr Val Asp Asn Ala Phe Gly Gly Gly Ser Ser  
 260 265 270  
 Asn Gln Met Phe Pro Leu Gln Asp Met Phe Met Tyr Met Gln Lys Pro  
 275 280 285  
 Tyr

<210> 105  
 <211> 1481  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (181)... (1188)

<400> 105  
 cgacccacgc gtccgagatt ctctcccagc tagctttctc aattcatttt tctttcttca 60  
 tctttcttctt gtgtgatctc tctttccaaa taagcttatac attcttacaa aaatatttct 120  
 gggtttctga tattgttctt gttctcttga atctttatta cttgaaaaac atataaagt 180  
 atg gcg gtt gtg gtt gaa gaa ggt gtg gtg ttg aat cat gga ggt gaa 228  
 Met Ala Val Val Val Glu Glu Gly Val Val Leu Asn His Gly Gly Glu  
 1 5 10 15  
 gag ctt gtg gat ttg cca cct ggt ttc agg ttt cat cca aca gac gaa 276  
 Glu Leu Val Asp Leu Pro Pro Gly Phe Arg Phe His Pro Thr Asp Glu  
 20 25 30  
 gag atc ata aca tgt tac ctt aag gag aag gtt tta aac agc cga ttc 324  
 Glu Ile Ile Thr Cys Tyr Leu Lys Glu Lys Val Leu Asn Ser Arg Phe  
 35 40 45  
 acg gct gtg gcc atg gga gaa gct gat ctc aac aag tgt gag cct tgg 372  
 Thr Ala Val Ala Met Gly Glu Ala Asp Leu Asn Lys Cys Glu Pro Trp  
 50 55 60  
 gat ttg cca aag agg gca aag atg ggg gag aaa gag ttc tac ttc ttc 420  
 Asp Leu Pro Lys Arg Ala Lys Met Gly Glu Lys Glu Phe Tyr Phe Phe  
 65 70 75 80  
 tgt caa agg gac agg aag tat ccg act ggg atg agg acg aac cgt gcg 468  
 Cys Gln Arg Asp Arg Lys Tyr Pro Thr Gly Met Arg Thr Asn Arg Ala  
 85 90 95  
 acg gag tca gga tac tgg aaa gcc acc ggg aag gat aag gag atc ttc 516  
 Thr Glu Ser Gly Tyr Trp Lys Ala Thr Gly Lys Asp Lys Glu Ile Phe  
 100 105 110  
 aaa ggc aaa ggt tgt ctc gtt ggg atg aag aaa aca ctt gtg ttt tat 564  
 Lys Gly Lys Gly Cys Leu Val Gly Met Lys Lys Thr Leu Val Phe Tyr

115	120	125	
aga gga aga gct cca aaa ggt gaa aag act aat tgg gtc atg cat gaa	612		
Arg Gly Arg Ala Pro Lys Gly Glu Lys Thr Asn Trp Val Met His Glu			
130	135	140	
tat cgt ctt gaa ggc aaa tat tcg tat tac aat ctc cca aaa tct gca	660		
Tyr Arg Leu Glu Gly Lys Tyr Ser Tyr Tyr Asn Leu Pro Lys Ser Ala			
145	150	155	160
agg gac gaa tgg gtc gtg tgt agg gtt ttt cac aag aac aat cct tct	708		
Arg Asp Glu Trp Val Val Cys Arg Val Phe His Lys Asn Asn Pro Ser			
165	170	175	
acc aca acc caa cca atg acg aga ata ccc gtt gaa gat ttc aca agg	756		
Thr Thr Thr Gln Pro Met Thr Arg Ile Pro Val Glu Asp Phe Thr Arg			
180	185	190	
atg gat tct cta gag aac att gat cat ctc cta gac ttc tca tct ctt	804		
Met Asp Ser Leu Glu Asn Ile Asp His Leu Leu Asp Phe Ser Ser Leu			
195	200	205	
cct cct ctc ata gac ccg agt ttc atg agt caa acc gaa caa cca aac	852		
Pro Pro Leu Ile Asp Pro Ser Phe Met Ser Gln Thr Glu Gln Pro Asn			
210	215	220	
ttc aaa ccc atc aac cct cca act tac gat atc tca tca cca atc caa	900		
Phe Lys Pro Ile Asn Pro Pro Thr Tyr Asp Ile Ser Ser Pro Ile Gln			
225	230	235	240
ccc cat cat ttc aat tct tac caa tca atc ttt aac cac cag gtt ttt	948		
Pro His His Phe Asn Ser Tyr Gln Ser Ile Phe Asn His Gln Val Phe			
245	250	255	
ggg tct gct tcg ggc tct acg tac aac aac aac aac gag atg atc aag	996		
Gly Ser Ala Ser Gly Ser Thr Tyr Asn Asn Asn Asn Glu Met Ile Lys			
260	265	270	
atg gag caa tca ctt gtt agt gta tct caa gaa aca tgc cta agc tca	1044		
Met Glu Gln Ser Leu Val Ser Val Ser Gln Glu Thr Cys Leu Ser Ser			
275	280	285	
gat gtg aac gcg aac atg act aca acc acg gag gta tct tcg ggt cct	1092		
Asp Val Asn Ala Asn Met Thr Thr Thr Thr Glu Val Ser Ser Gly Pro			
290	295	300	
gta atg aaa caa gaa atg ggg atg atg gga atg gtg aat ggt agc aag	1140		
Val Met Lys Gln Glu Met Gly Met Met Gly Met Val Asn Gly Ser Lys			
305	310	315	320
tcg tat gaa gat cta tgt gac ttg agg ggg gac ttg tgg gac ttc taa	1188		
Ser Tyr Glu Asp Leu Cys Asp Leu Arg Gly Asp Leu Trp Asp Phe *			
325	330	335	
ttaatcattt gactgtggtg aaagagtata ttgttggtgga tttaaatacat gttagttaat	1248		
acatacat ataggattta ctaggagctt aatcctagtt aactattttc acttcattga	1308		
tattatttaa ttagttgatt gtttaattag ttatacttt atagtgtggt taaaaaagaa	1368		



aagaaaggat tgtgataatt tgggatttta gtgcataagt tatatctcaa tgtaaaactgt 1428  
 atttgtatcc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 1481

<210> 106

<211> 335

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (21)...(149)

<223> Conserved domain

<400> 106

Met	Ala	Val	Val	Val	Glu	Glu	Gly	Val	Val	Leu	Asn	His	Gly	Gly	Glu
1				5					10					15	
Glu	Leu	Val	Asp	Leu	Pro	Pro	Gly	Phe	Arg	Phe	His	Pro	Thr	Asp	Glu
			20					25					30		
Glu	Ile	Ile	Thr	Cys	Tyr	Leu	Lys	Glu	Lys	Val	Leu	Asn	Ser	Arg	Phe
		35					40					45			
Thr	Ala	Val	Ala	Met	Gly	Glu	Ala	Asp	Leu	Asn	Lys	Cys	Glu	Pro	Trp
		50				55					60				
Asp	Leu	Pro	Lys	Arg	Ala	Lys	Met	Gly	Glu	Lys	Glu	Phe	Tyr	Phe	Phe
				70						75					80
Cys	Gln	Arg	Asp	Arg	Lys	Tyr	Pro	Thr	Gly	Met	Arg	Thr	Asn	Arg	Ala
				85					90					95	
Thr	Glu	Ser	Gly	Tyr	Trp	Lys	Ala	Thr	Gly	Lys	Asp	Lys	Glu	Ile	Phe
			100					105					110		
Lys	Gly	Lys	Gly	Cys	Leu	Val	Gly	Met	Lys	Lys	Thr	Leu	Val	Phe	Tyr
		115					120					125			
Arg	Gly	Arg	Ala	Pro	Lys	Gly	Glu	Lys	Thr	Asn	Trp	Val	Met	His	Glu
		130				135					140				
Tyr	Arg	Leu	Glu	Gly	Lys	Tyr	Ser	Tyr	Tyr	Asn	Leu	Pro	Lys	Ser	Ala
				150						155					160
Arg	Asp	Glu	Trp	Val	Val	Cys	Arg	Val	Phe	His	Lys	Asn	Asn	Pro	Ser
				165					170					175	
Thr	Thr	Thr	Gln	Pro	Met	Thr	Arg	Ile	Pro	Val	Glu	Asp	Phe	Thr	Arg
			180					185					190		
Met	Asp	Ser	Leu	Glu	Asn	Ile	Asp	His	Leu	Leu	Asp	Phe	Ser	Ser	Leu
		195					200					205			
Pro	Pro	Leu	Ile	Asp	Pro	Ser	Phe	Met	Ser	Gln	Thr	Glu	Gln	Pro	Asn
		210				215					220				
Phe	Lys	Pro	Ile	Asn	Pro	Pro	Thr	Tyr	Asp	Ile	Ser	Ser	Pro	Ile	Gln
				230					235						240
Pro	His	His	Phe	Asn	Ser	Tyr	Gln	Ser	Ile	Phe	Asn	His	Gln	Val	Phe
				245					250					255	
Gly	Ser	Ala	Ser	Gly	Ser	Thr	Tyr	Asn	Asn	Asn	Asn	Glu	Met	Ile	Lys
			260					265					270		
Met	Glu	Gln	Ser	Leu	Val	Ser	Val	Ser	Gln	Glu	Thr	Cys	Leu	Ser	Ser
		275					280					285			
Asp	Val	Asn	Ala	Asn	Met	Thr	Thr	Thr	Thr	Glu	Val	Ser	Ser	Gly	Pro
		290				295					300				
Val	Met	Lys	Gln	Glu	Met	Gly	Met	Met	Gly	Met	Val	Asn	Gly	Ser	Lys
				310					315						320
Ser	Tyr	Glu	Asp	Leu	Cys	Asp	Leu	Arg	Gly	Asp	Leu	Trp	Asp	Phe	
				325				330						335	

<210> 107  
 <211> 1130  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (116)...(1024)

<221> misc\_feature  
 <222> 1100  
 <223> n = A,T,C or G

<400> 107  
 ccaagatcga ctcttacttc gaatctctct caactttctt cctcagctta cgggaacttc 60  
 cacacatata catccacaag aacccatata gaagattcat cctacatata ttac atg 118  
 Met  
 1

gat cag tac tca tcc tct ttg gtc gat act tca tta gat ctc act att 166  
 Asp Gln Tyr Ser Ser Ser Leu Val Asp Thr Ser Leu Asp Leu Thr Ile  
 5 10 15

ggc gtt act cgt atg cga gtt gaa gaa gat cca ccg aca agt gct ttg 214  
 Gly Val Thr Arg Met Arg Val Glu Glu Asp Pro Pro Thr Ser Ala Leu  
 20 25 30

gtg gaa gaa tta aac cga gtt agt gct gag aac aag aag ctc tcg gag 262  
 Val Glu Glu Leu Asn Arg Val Ser Ala Glu Asn Lys Lys Leu Ser Glu  
 35 40 45

atg cta act ttg atg tgt gac aac tac aac gtc ttg agg aag caa ctt 310  
 Met Leu Thr Leu Met Cys Asp Asn Tyr Asn Val Leu Arg Lys Gln Leu  
 50 55 60 65

atg gaa tat gtt aac aag agc aac ata acc gag agg gat caa atc agc 358  
 Met Glu Tyr Val Asn Lys Ser Asn Ile Thr Glu Arg Asp Gln Ile Ser  
 70 75 80

cct ccc aag aaa cgc aaa tcc ccg gcg aga gag gac gca ttc agc tgc 406  
 Pro Pro Lys Lys Arg Lys Ser Pro Ala Arg Glu Asp Ala Phe Ser Cys  
 85 90 95

gcg gtt att ggc gga gtg tcg gag agt agc tca acg gat caa gat gag 454  
 Ala Val Ile Gly Gly Val Ser Glu Ser Ser Ser Thr Asp Gln Asp Glu  
 100 105 110

tat ttg tgt aag aag cag aga gaa gag act gtc gtg aag gag aaa gtc 502  
 Tyr Leu Cys Lys Lys Gln Arg Glu Glu Thr Val Val Lys Glu Lys Val  
 115 120 125

tca agg gtc tat tac aag acc gaa gct tct gac act acc ctc gtt gtg 550  
 Ser Arg Val Tyr Tyr Lys Thr Glu Ala Ser Asp Thr Thr Leu Val Val  
 130 135 140 145

aaa gat ggg tat caa tgg agg aaa tat gga cag aaa gtg act aga gac 598

Lys Asp Gly Tyr Gln Trp Arg Lys Tyr Gly Gln Lys Val Thr Arg Asp  
 150 155 160  
 aat cca tct cca aga gct tac ttc aaa tgt gct tgt gct cca agc tgt 646  
 Asn Pro Ser Pro Arg Ala Tyr Phe Lys Cys Ala Cys Ala Pro Ser Cys  
 165 170 175  
 tct gtc aaa aag aag gtt cag aga agt gtg gag gat cag tcc gtg tta 694  
 Ser Val Lys Lys Lys Val Gln Arg Ser Val Glu Asp Gln Ser Val Leu  
 180 185 190  
 gtt gca act tat gag ggt gaa cac aac cat cca atg cca tcg cag atc 742  
 Val Ala Thr Tyr Glu Gly His Asn His Pro Met Pro Ser Gln Ile  
 195 200 205  
 gat tca aac aat ggc tta aac cgc cac atc tct cat ggt ggt tca gct 790  
 Asp Ser Asn Asn Gly Leu Asn Arg His Ile Ser His Gly Gly Ser Ala  
 210 215 220 225  
 tca aca ccc gtt gca gca aac aga aga agt agc ttg act gtg ccg gtg 838  
 Ser Thr Pro Val Ala Ala Asn Arg Arg Ser Ser Leu Thr Val Pro Val  
 230 235 240  
 act acc gta gat atg att gaa tcg aag aaa gtg acg agc cca acg tca 886  
 Thr Thr Val Asp Met Ile Glu Ser Lys Lys Val Thr Ser Pro Thr Ser  
 245 250 255  
 aga atc gat ttt ccc caa gtt cag aaa ctt ttg gtg gag caa atg gct 934  
 Arg Ile Asp Phe Pro Gln Val Gln Lys Leu Leu Val Glu Gln Met Ala  
 260 265 270  
 tct tcc tta acc aaa gat cct aac ttt aca gca gct tta gca gca gct 982  
 Ser Ser Leu Thr Lys Asp Pro Asn Phe Thr Ala Ala Leu Ala Ala Ala  
 275 280 285  
 gtt acc gga aaa ttg tat caa cag aat cat acc gag aaa tag 1024  
 Val Thr Gly Lys Leu Tyr Gln Gln Asn His Thr Glu Lys \*  
 290 295 300  
 ttttagcttca aattccgtta gagtttttag atttgaattt gtcattgagta agagaaagag 1084  
 agtagattat aatccnttgt gatactgaaa aaaaaaaaaa aaaaaa 1130  
 <210> 108  
 <211> 302  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <221> DOMAIN  
 <222> (146)...(203)  
 <223> Conserved domain  
 <400> 108  
 Met Asp Gln Tyr Ser Ser Ser Leu Val Asp Thr Ser Leu Asp Leu Thr  
 1 5 10 15  
 Ile Gly Val Thr Arg Met Arg Val Glu Glu Asp Pro Pro Thr Ser Ala  
 20 25 30

Leu Val Glu Glu Leu Asn Arg Val Ser Ala Glu Asn Lys Lys Leu Ser  
           35                          40                          45  
 Glu Met Leu Thr Leu Met Cys Asp Asn Tyr Asn Val Leu Arg Lys Gln  
           50                          55                          60  
 Leu Met Glu Tyr Val Asn Lys Ser Asn Ile Thr Glu Arg Asp Gln Ile  
           65                          70                          75                          80  
 Ser Pro Pro Lys Lys Arg Lys Ser Pro Ala Arg Glu Asp Ala Phe Ser  
                           85                          90                          95  
 Cys Ala Val Ile Gly Gly Val Ser Glu Ser Ser Ser Thr Asp Gln Asp  
                           100                          105                          110  
 Glu Tyr Leu Cys Lys Lys Gln Arg Glu Glu Thr Val Val Lys Glu Lys  
                           115                          120                          125  
 Val Ser Arg Val Tyr Tyr Lys Thr Glu Ala Ser Asp Thr Thr Leu Val  
           130                          135                          140  
 Val Lys Asp Gly Tyr Gln Trp Arg Lys Tyr Gly Gln Lys Val Thr Arg  
           145                          150                          155                          160  
 Asp Asn Pro Ser Pro Arg Ala Tyr Phe Lys Cys Ala Cys Ala Pro Ser  
                           165                          170                          175  
 Cys Ser Val Lys Lys Lys Val Gln Arg Ser Val Glu Asp Gln Ser Val  
                           180                          185                          190  
 Leu Val Ala Thr Tyr Glu Gly Glu His Asn His Pro Met Pro Ser Gln  
           195                          200                          205  
 Ile Asp Ser Asn Asn Gly Leu Asn Arg His Ile Ser His Gly Gly Ser  
           210                          215                          220  
 Ala Ser Thr Pro Val Ala Ala Asn Arg Arg Ser Ser Leu Thr Val Pro  
           225                          230                          235                          240  
 Val Thr Thr Val Asp Met Ile Glu Ser Lys Lys Val Thr Ser Pro Thr  
                           245                          250                          255  
 Ser Arg Ile Asp Phe Pro Gln Val Gln Lys Leu Leu Val Glu Gln Met  
                           260                          265                          270  
 Ala Ser Ser Leu Thr Lys Asp Pro Asn Phe Thr Ala Ala Leu Ala Ala  
           275                          280                          285  
 Ala Val Thr Gly Lys Leu Tyr Gln Gln Asn His Thr Glu Lys  
           290                          295                          300

&lt;210&gt; 109

&lt;211&gt; 1449

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) ... (1449)

&lt;400&gt; 109

atg gtg gct atg ttt caa gaa gat aat gga aca tct tct gta gct tca 48  
 Met Val Ala Met Phe Gln Glu Asp Asn Gly Thr Ser Ser Val Ala Ser  
   1                          5                          10                          15

tca cca ctt caa gtc ttc tca act atg tca ctc aac aga ccg act ctc 96  
 Ser Pro Leu Gln Val Phe Ser Thr Met Ser Leu Asn Arg Pro Thr Leu  
           20                          25                          30

ctc gct tct tca tct ccg ttt cat tgt ctc aaa gat ctc aaa cca gag 144  
 Leu Ala Ser Ser Ser Pro Phe His Cys Leu Lys Asp Leu Lys Pro Glu  
           35                          40                          45

gag cgt ggt ctc tac tta atc cac ctc ttg cta act tgt gcc aac cac	192
Glu Arg Gly Leu Tyr Leu Ile His Leu Leu Leu Thr Cys Ala Asn His	
50 55 60	
gtg gct tca ggt agc ctc caa aac gct aac gca gcg ctc gag cag ctc	240
Val Ala Ser Gly Ser Leu Gln Asn Ala Asn Ala Ala Leu Glu Gln Leu	
65 70 75 80	
tct cac ctc gct tct cct gac ggc gac acg atg cag cga atc gct gct	288
Ser His Leu Ala Ser Pro Asp Gly Asp Thr Met Gln Arg Ile Ala Ala	
85 90 95	
tac ttc acc gaa gcg ctt gct aac aga atc ctt aag tcc tgg cct ggt	336
Tyr Phe Thr Glu Ala Leu Ala Asn Arg Ile Leu Lys Ser Trp Pro Gly	
100 105 110	
ctt tac aag gct ctt aac gca act cag aca aga act aac aat gtc tct	384
Leu Tyr Lys Ala Leu Asn Ala Thr Gln Thr Arg Thr Asn Asn Val Ser	
115 120 125	
gag gag att cat gtt aga aga ctc ttc ttt gag atg ttc ccg ata ctc	432
Glu Glu Ile His Val Arg Arg Leu Phe Phe Glu Met Phe Pro Ile Leu	
130 135 140	
aaa gtc tct tac ttg ctc act aat cga gct ata ctc gag gct atg gaa	480
Lys Val Ser Tyr Leu Leu Thr Asn Arg Ala Ile Leu Glu Ala Met Glu	
145 150 155 160	
gga gag aag atg gtt cat gtg att gat ctc gat gct tct gag cca gct	528
Gly Glu Lys Met Val His Val Ile Asp Leu Asp Ala Ser Glu Pro Ala	
165 170 175	
caa tgg ctt gct ttg ctt caa gct ttt aac tct agg cct gaa ggt cca	576
Gln Trp Leu Ala Leu Leu Gln Ala Phe Asn Ser Arg Pro Glu Gly Pro	
180 185 190	
cct cat ttg aga atc act ggt gtt cat cac cag aag gaa gtg ctt gaa	624
Pro His Leu Arg Ile Thr Gly Val His His Gln Lys Glu Val Leu Glu	
195 200 205	
caa atg gct cat aga ctc att gag gaa gca gag aaa ctc gat atc ccg	672
Gln Met Ala His Arg Leu Ile Glu Glu Ala Glu Lys Leu Asp Ile Pro	
210 215 220	
ttt cag ttt aat ccc gtt gtg agt agg tta gac tgt tta aat gta gaa	720
Phe Gln Phe Asn Pro Val Val Ser Arg Leu Asp Cys Leu Asn Val Glu	
225 230 235 240	
cag ttg cgg gtt aaa aca gga gag gcc tta gcc gtt agc tcg gtt ctt	768
Gln Leu Arg Val Lys Thr Gly Glu Ala Leu Ala Val Ser Ser Val Leu	
245 250 255	
caa ttg cat acc ttc ttg gcc tct gat gat gat ctc atg aga aag aac	816
Gln Leu His Thr Phe Leu Ala Ser Asp Asp Asp Leu Met Arg Lys Asn	
260 265 270	

tgc gct tta cgg ttt cag aac aac cct agt gga gtt gac ttg cag aga	864
Cys Ala Leu Arg Phe Gln Asn Asn Pro Ser Gly Val Asp Leu Gln Arg	
275 280 285	
ggt cta atg atg agc cat ggc tct gca gct gag gca cgt gag aat gat	912
Val Leu Met Met Ser His Gly Ser Ala Ala Glu Ala Arg Glu Asn Asp	
290 295 300	
atg agt aac aac aat ggg tat agc cct agc ggt gac tcg gcc tca tct	960
Met Ser Asn Asn Asn Gly Tyr Ser Pro Ser Gly Asp Ser Ala Ser Ser	
305 310 315 320	
ttg cct tta cca agt tca gga agg act gat agc ttc ctc aat gct att	1008
Leu Pro Leu Pro Ser Ser Gly Arg Thr Asp Ser Phe Leu Asn Ala Ile	
325 330 335	
tgg ggt ttg tct cca aag gtc atg gtg gtc act gag caa gac tca gac	1056
Trp Gly Leu Ser Pro Lys Val Met Val Val Thr Glu Gln Asp Ser Asp	
340 345 350	
cac aac ggc tcc aca cta atg gag agg cta tta gaa tca ctt tac acc	1104
His Asn Gly Ser Thr Leu Met Glu Arg Leu Leu Glu Ser Leu Tyr Thr	
355 360 365	
tac gca gca ttg ttt gat tgc ttg gaa aca aaa gtt cca aga acg tct	1152
Tyr Ala Ala Leu Phe Asp Cys Leu Glu Thr Lys Val Pro Arg Thr Ser	
370 375 380	
caa gat agg atc aaa gtg gag aag atg ctc ttc ggg gag gag atc aag	1200
Gln Asp Arg Ile Lys Val Glu Lys Met Leu Phe Gly Glu Glu Ile Lys	
385 390 395 400	
aac atc ata tcc tgc gag gga ttt gag aga aga gaa aga cac gag aag	1248
Asn Ile Ile Ser Cys Glu Gly Phe Glu Arg Arg Glu Arg His Glu Lys	
405 410 415	
ctt gag aaa tgg agc cag agg atc gat ttg gct ggt ttt ggg aat gtt	1296
Leu Glu Lys Trp Ser Gln Arg Ile Asp Leu Ala Gly Phe Gly Asn Val	
420 425 430	
cct ctt agc tat tat gcg atg ttg cag gct agg aga ttg ctt caa ggg	1344
Pro Leu Ser Tyr Tyr Ala Met Leu Gln Ala Arg Arg Leu Leu Gln Gly	
435 440 445	
tgc ggt ttt gat ggg tat aga atc aag gaa gag agc ggg tgc gca gta	1392
Cys Gly Phe Asp Gly Tyr Arg Ile Lys Glu Glu Ser Gly Cys Ala Val	
450 455 460	
att tgc tgg caa gat cga cct cta tac tcg gta tca gct tgg aga tgc	1440
Ile Cys Trp Gln Asp Arg Pro Leu Tyr Ser Val Ser Ala Trp Arg Cys	
465 470 475 480	
agg aag tga	1449
Arg Lys *	

<210> 110  
 <211> 482  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (225)...(242)  
 <223> Conserved domain

<400> 110  
 Met Val Ala Met Phe Gln Glu Asp Asn Gly Thr Ser Ser Val Ala Ser  
 1 5 10 15  
 Ser Pro Leu Gln Val Phe Ser Thr Met Ser Leu Asn Arg Pro Thr Leu  
 20 25 30  
 Leu Ala Ser Ser Ser Pro Phe His Cys Leu Lys Asp Leu Lys Pro Glu  
 35 40 45  
 Glu Arg Gly Leu Tyr Leu Ile His Leu Leu Leu Thr Cys Ala Asn His  
 50 55 60  
 Val Ala Ser Gly Ser Leu Gln Asn Ala Asn Ala Ala Leu Glu Gln Leu  
 65 70 75 80  
 Ser His Leu Ala Ser Pro Asp Gly Asp Thr Met Gln Arg Ile Ala Ala  
 85 90 95  
 Tyr Phe Thr Glu Ala Leu Ala Asn Arg Ile Leu Lys Ser Trp Pro Gly  
 100 105 110  
 Leu Tyr Lys Ala Leu Asn Ala Thr Gln Thr Arg Thr Asn Asn Val Ser  
 115 120 125  
 Glu Glu Ile His Val Arg Arg Leu Phe Phe Glu Met Phe Pro Ile Leu  
 130 135 140  
 Lys Val Ser Tyr Leu Leu Thr Asn Arg Ala Ile Leu Glu Ala Met Glu  
 145 150 155 160  
 Gly Glu Lys Met Val His Val Ile Asp Leu Asp Ala Ser Glu Pro Ala  
 165 170 175  
 Gln Trp Leu Ala Leu Leu Gln Ala Phe Asn Ser Arg Pro Glu Gly Pro  
 180 185 190  
 Pro His Leu Arg Ile Thr Gly Val His His Gln Lys Glu Val Leu Glu  
 195 200 205  
 Gln Met Ala His Arg Leu Ile Glu Glu Ala Glu Lys Leu Asp Ile Pro  
 210 215 220  
 Phe Gln Phe Asn Pro Val Val Ser Arg Leu Asp Cys Leu Asn Val Glu  
 225 230 235 240  
 Gln Leu Arg Val Lys Thr Gly Glu Ala Leu Ala Val Ser Ser Val Leu  
 245 250 255  
 Gln Leu His Thr Phe Leu Ala Ser Asp Asp Asp Leu Met Arg Lys Asn  
 260 265 270  
 Cys Ala Leu Arg Phe Gln Asn Asn Pro Ser Gly Val Asp Leu Gln Arg  
 275 280 285  
 Val Leu Met Met Ser His Gly Ser Ala Ala Glu Ala Arg Glu Asn Asp  
 290 295 300  
 Met Ser Asn Asn Asn Gly Tyr Ser Pro Ser Gly Asp Ser Ala Ser Ser  
 305 310 315 320  
 Leu Pro Leu Pro Ser Ser Gly Arg Thr Asp Ser Phe Leu Asn Ala Ile  
 325 330 335  
 Trp Gly Leu Ser Pro Lys Val Met Val Val Thr Glu Gln Asp Ser Asp  
 340 345 350  
 His Asn Gly Ser Thr Leu Met Glu Arg Leu Leu Glu Ser Leu Tyr Thr  
 355 360 365

Tyr Ala Ala Leu Phe Asp Cys Leu Glu Thr Lys Val Pro Arg Thr Ser  
 370 375 380  
 Gln Asp Arg Ile Lys Val Glu Lys Met Leu Phe Gly Glu Glu Ile Lys  
 385 390 395 400  
 Asn Ile Ile Ser Cys Glu Gly Phe Glu Arg Arg Glu Arg His Glu Lys  
 405 410 415  
 Leu Glu Lys Trp Ser Gln Arg Ile Asp Leu Ala Gly Phe Gly Asn Val  
 420 425 430  
 Pro Leu Ser Tyr Tyr Ala Met Leu Gln Ala Arg Arg Leu Leu Gln Gly  
 435 440 445  
 Cys Gly Phe Asp Gly Tyr Arg Ile Lys Glu Glu Ser Gly Cys Ala Val  
 450 455 460  
 Ile Cys Trp Gln Asp Arg Pro Leu Tyr Ser Val Ser Ala Trp Arg Cys  
 465 470 475 480  
 Arg Lys

<210> 111  
 <211> 1388  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (280)...(1095)

<400> 111  
 ccaaaatcta ggggttttctt ctgcgccaat ttcacttttc ttctacgaaa ttctccattc 60  
 ctgccggctg tcgggttttc tgaatcgatt ctccctcacc aacttcttct ctggttctgt 120  
 tcgattctga ttttttttca aggtcaattt tttcttctct ttaaactctg caaaatcgtg 180  
 atcgattaaa ttcacctcag gggtttttga tttctgaaag aagttaatct tcttgaagg 240  
 cgattgcaaa agagtgtctt gctgtgaatt tccactgag atg caa tca aaa ccg 294  
 Met Gln Ser Lys Pro  
 1 5  
 gga aga gaa aac gaa gag gaa gtc aat aat cac cat gct gtt cag cag 342  
 Gly Arg Glu Asn Glu Glu Val Asn Asn His His Ala Val Gln Gln  
 10 15 20  
 ccg atg atg tat gca gag ccc tgg tgg aaa aac aac tcc ttt ggt gtt 390  
 Pro Met Met Tyr Ala Glu Pro Trp Trp Lys Asn Asn Ser Phe Gly Val  
 25 30 35  
 gta cct caa gcg aga cct tct gga att cca tca aat tcc tct tct ttg 438  
 Val Pro Gln Ala Arg Pro Ser Gly Ile Pro Ser Asn Ser Ser Ser Leu  
 40 45 50  
 gat tgc ccc aat ggt tcc gag tca aac gat gtt cat tca gca tct gaa 486  
 Asp Cys Pro Asn Gly Ser Glu Ser Asn Asp Val His Ser Ala Ser Glu  
 55 60 65  
 gac ggt gcg ttg aat ggt gaa aac gat ggc act tgg aag gat tca caa 534  
 Asp Gly Ala Leu Asn Gly Glu Asn Asp Gly Thr Trp Lys Asp Ser Gln  
 70 75 80 85  
 gct gca act tcc tct cgt tca gat aat cac gga atg gaa gga aat gac 582



Ala Ala Thr Ser Ser Arg Ser Asp Asn His Gly Met Glu Gly Asn Asp  
90 95 100

cca gcg ctc tct atc cgt aac atg cat gat cag cca ctt gta caa cca 630  
Pro Ala Leu Ser Ile Arg Asn Met His Asp Gln Pro Leu Val Gln Pro  
105 110 115

cca gag ctt gtt gga cac tat atc gct tgt gtc cca aac cca tat cag 678  
Pro Glu Leu Val Gly His Tyr Ile Ala Cys Val Pro Asn Pro Tyr Gln  
120 125 130

gat cca tat tat ggg gga ttg atg gga gca tat ggt cat cag caa ttg 726  
Asp Pro Tyr Tyr Gly Gly Leu Met Gly Ala Tyr Gly His Gln Gln Leu  
135 140 145

ggt ttt cgt cca tat ctt gga atg cct cgt gaa aga aca gct ctg cca 774  
Gly Phe Arg Pro Tyr Leu Gly Met Pro Arg Glu Arg Thr Ala Leu Pro  
150 155 160 165

ctt gac atg gca caa gag ccc gtt tat gtg aat gca aag cag tac gag 822  
Leu Asp Met Ala Gln Glu Pro Val Tyr Val Asn Ala Lys Gln Tyr Glu  
170 175 180

gga att cta agg cga aga aaa gca cgt gcc aag gca gag cta gag agg 870  
Gly Ile Leu Arg Arg Arg Lys Ala Arg Ala Lys Ala Glu Leu Glu Arg  
185 190 195

aaa gtc atc cgg gac aga aag cca tat ctt cac gag tca aga cac aag 918  
Lys Val Ile Arg Asp Arg Lys Pro Tyr Leu His Glu Ser Arg His Lys  
200 205 210

cat gca atg aga agg gca cga gcg agt gga ggc cgg ttt gcg aag aaa 966  
His Ala Met Arg Arg Ala Arg Ala Ser Gly Gly Arg Phe Ala Lys Lys  
215 220 225

agt gag gta gaa gcg gga gag gat gca gga ggg aga gac aga gaa agg 1014  
Ser Glu Val Glu Ala Gly Glu Asp Ala Gly Gly Arg Asp Arg Glu Arg  
230 235 240 245

ggt tca gca acc aac tca tca ggc tct gaa caa gtt gag aca gac tct 1062  
Gly Ser Ala Thr Asn Ser Ser Gly Ser Glu Gln Val Glu Thr Asp Ser  
250 255 260

aat gag acc ctg aat tct tct ggt gca cca taa taaaaaaagc caaagctctg 1115  
Asn Glu Thr Leu Asn Ser Ser Gly Ala Pro \*

265 270

agaggagaga gagacacaca ctttggttaa tataatccat tgcctcaaac cggcaaatca 1175  
ttcttggtt ttctgttttt gtgtttgcta gttgttcttg tcagagtctc atattgtgtg 1235  
ggtttaacag ttatgatgaa tgtacaaaga gcgagttatg ttaggtgtta gattttggag 1295  
acaagagaca aaggaatagc aagtaggtct tgtttttatt ctttgacctt tttttctct 1355  
tttgcaaaat tgaaaaatac gtttgcttaa aaa 1388

<210> 112  
<211> 271  
<212> PRT  
<213> Arabidopsis thaliana

&lt;400&gt; 112

```

Met Gln Ser Lys Pro Gly Arg Glu Asn Glu Glu Glu Val Asn Asn His
 1          5          10          15
His Ala Val Gln Gln Pro Met Met Tyr Ala Glu Pro Trp Trp Lys Asn
          20          25          30
Asn Ser Phe Gly Val Val Pro Gln Ala Arg Pro Ser Gly Ile Pro Ser
          35          40          45
Asn Ser Ser Ser Leu Asp Cys Pro Asn Gly Ser Glu Ser Asn Asp Val
          50          55          60
His Ser Ala Ser Glu Asp Gly Ala Leu Asn Gly Glu Asn Asp Gly Thr
          65          70          75          80
Trp Lys Asp Ser Gln Ala Ala Thr Ser Ser Arg Ser Asp Asn His Gly
          85          90          95
Met Glu Gly Asn Asp Pro Ala Leu Ser Ile Arg Asn Met His Asp Gln
          100          105          110
Pro Leu Val Gln Pro Pro Glu Leu Val Gly His Tyr Ile Ala Cys Val
          115          120          125
Pro Asn Pro Tyr Gln Asp Pro Tyr Tyr Gly Gly Leu Met Gly Ala Tyr
          130          135          140
Gly His Gln Gln Leu Gly Phe Arg Pro Tyr Leu Gly Met Pro Arg Glu
          145          150          155          160
Arg Thr Ala Leu Pro Leu Asp Met Ala Gln Glu Pro Val Tyr Val Asn
          165          170          175
Ala Lys Gln Tyr Glu Gly Ile Leu Arg Arg Arg Lys Ala Arg Ala Lys
          180          185          190
Ala Glu Leu Glu Arg Lys Val Ile Arg Asp Arg Lys Pro Tyr Leu His
          195          200          205
Glu Ser Arg His Lys His Ala Met Arg Arg Ala Arg Ala Ser Gly Gly
          210          215          220
Arg Phe Ala Lys Lys Ser Glu Val Glu Ala Gly Glu Asp Ala Gly Gly
          225          230          235          240
Arg Asp Arg Glu Arg Gly Ser Ala Thr Asn Ser Ser Gly Ser Glu Gln
          245          250          255
Val Glu Thr Asp Ser Asn Glu Thr Leu Asn Ser Ser Gly Ala Pro
          260          265          270

```

&lt;210&gt; 113

&lt;211&gt; 832

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (103)...(528)

&lt;400&gt; 113

```

gagcggttcg tagaaaaatt cgatttctct aaagccctaa aactaaaacg actatcccca 60
attccaagtt ctagggtttc catcttcccc aatctagtat aa atg gcg gat acg 114
                               Met Ala Asp Thr
                               1

```

```

cct tcg agc cca gct gga gat ggc gga gaa agc ggc ggt tcc gtt agg 162
Pro Ser Ser Pro Ala Gly Asp Gly Gly Glu Ser Gly Gly Ser Val Arg
 5          10          15          20

```

gag cag gat cga tac ctt cct ata gct aat atc agc agg atc atg aag 210  
 Glu Gln Asp Arg Tyr Leu Pro Ile Ala Asn Ile Ser Arg Ile Met Lys  
                   25                                  30                                  35

aaa gcg ttg cct cct aat ggt aag att gga aaa gat gct aag gat aca 258  
 Lys Ala Leu Pro Pro Asn Gly Lys Ile Gly Lys Asp Ala Lys Asp Thr  
                   40                                  45                                  50

gtt cag gaa tgc gtc tct gag ttc atc agc ttc atc act agc gag gcc 306  
 Val Gln Glu Cys Val Ser Glu Phe Ile Ser Phe Ile Thr Ser Glu Ala  
                   55                                  60                                  65

agt gat aag tgt caa aaa gag aaa agg aaa act gtg aat ggt gat gat 354  
 Ser Asp Lys Cys Gln Lys Glu Lys Arg Lys Thr Val Asn Gly Asp Asp  
                   70                                  75                                  80

ttg ttg tgg gca atg gca aca tta gga ttt gag gat tac ctg gaa cct 402  
 Leu Leu Trp Ala Met Ala Thr Leu Gly Phe Glu Asp Tyr Leu Glu Pro  
                   85                                  90                                  95                                  100

cta aag ata tac cta gcg agg tac agg gag ttg gag ggt gat aat aag 450  
 Leu Lys Ile Tyr Leu Ala Arg Tyr Arg Glu Leu Glu Gly Asp Asn Lys  
                   105                                  110                                  115

gga tca gga aag agt gga gat gga tca aat aga gat gct ggt ggc ggt 498  
 Gly Ser Gly Lys Ser Gly Asp Gly Ser Asn Arg Asp Ala Gly Gly Gly  
                   120                                  125                                  130

gtt tct ggt gaa gaa atg ccg agc tgg taa aagaagttgc aagtagtgat 548  
 Val Ser Gly Glu Glu Met Pro Ser Trp \*  
                   135                                  140

taagaacaat cgccaaatga tcaagggaaa ttagagatca gtgagttggt tatagttgag 608  
 ctgatcgaca actatcttcgg gtttactctc aatttcggtt atgttagttt gaacggtttg 668  
 tttattgttt ccggtttagt tggttgtatt taaagatttc tctgtagat gttgagaaca 728  
 cttgaatgaa ggaaaaattt gtccacatcc tgttggtatt ttcgattcac tttcggaatt 788  
 tcatagctaa tttattctca ttttaataacca aatccttaaa ttaa 832

<210> 114

<211> 141

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (20)...(109)

<223> Conserved domain

<400> 114

Met Ala Asp Thr Pro Ser Ser Pro Ala Gly Asp Gly Gly Glu Ser Gly  
   1                                  5                                  10                                  15  
 Gly Ser Val Arg Glu Gln Asp Arg Tyr Leu Pro Ile Ala Asn Ile Ser  
                   20                                  25                                  30  
 Arg Ile Met Lys Lys Ala Leu Pro Asn Gly Lys Ile Gly Lys Asp  
                   35                                  40                                  45  
 Ala Lys Asp Thr Val Gln Glu Cys Val Ser Glu Phe Ile Ser Phe Ile  
                   50                                  55                                  60

Thr Ser Glu Ala Ser Asp Lys Cys Gln Lys Glu Lys Arg Lys Thr Val  
 65 70 75 80  
 Asn Gly Asp Asp Leu Leu Trp Ala Met Ala Thr Leu Gly Phe Glu Asp  
 85 90 95  
 Tyr Leu Glu Pro Leu Lys Ile Tyr Leu Ala Arg Tyr Arg Glu Leu Glu  
 100 105 110  
 Gly Asp Asn Lys Gly Ser Gly Lys Ser Gly Asp Gly Ser Asn Arg Asp  
 115 120 125  
 Ala Gly Gly Gly Val Ser Gly Glu Glu Met Pro Ser Trp  
 130 135 140

<210> 115

<211> 1662

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)...(1662)

<400> 115

atg ggt gat aca gag aag tgt aac agt gat atg atc cag aga ctt cat 48  
 Met Gly Asp Thr Glu Lys Cys Asn Ser Asp Met Ile Gln Arg Leu His  
 1 5 10 15

tca tct ttc ggc act act tct tct tcc att ccc aaa aat ccc att tct 96  
 Ser Ser Phe Gly Thr Thr Ser Ser Ser Ile Pro Lys Asn Pro Ile Ser  
 20 25 30

cag ctc gat tta aac cct aat ttc atc cgc tca tca gct cct caa ttc 144  
 Gln Leu Asp Leu Asn Pro Asn Phe Ile Arg Ser Ser Ala Pro Gln Phe  
 35 40 45

tcc aag cct ttc agt gac agt ggc aaa cga atc ggt gtt cct ccg tcg 192  
 Ser Lys Pro Phe Ser Asp Ser Gly Lys Arg Ile Gly Val Pro Pro Ser  
 50 55 60

cac ccc aac tta atc cca ccg act tct ccg ttt tct cag atc ccg acc 240  
 His Pro Asn Leu Ile Pro Pro Thr Ser Pro Phe Ser Gln Ile Pro Thr  
 65 70 75 80

acc cga caa ccc ggt tcg cat aat ttt aac ccg gga gga gct aat cat 288  
 Thr Arg Gln Pro Gly Ser His Asn Phe Asn Pro Gly Gly Ala Asn His  
 85 90 95

tca cgg tca atg tca cag ccc aac tct ttc ttc tct ttt gac tcc tta 336  
 Ser Arg Ser Met Ser Gln Pro Asn Ser Phe Phe Ser Phe Asp Ser Leu  
 100 105 110

cct ccg tta agc cct tct ccg ttt cga gat cac gat gtt tca atg gag 384  
 Pro Pro Leu Ser Pro Ser Pro Phe Arg Asp His Asp Val Ser Met Glu  
 115 120 125

gat aga gat tcc ggc gtg ttt aac agc aac cat tcg ttg cct cca tcg 432  
 Asp Arg Asp Ser Gly Val Phe Asn Ser Asn His Ser Leu Pro Pro Ser  
 130 135 140

cgc	ttc	acg	agg	tgt	aat	tcg	acc	tct	tct	agc	tcc	ttg	aga	gtc	ggc	480
Pro	Phe	Thr	Arg	Cys	Asn	Ser	Thr	Ser	Ser	Ser	Ser	Leu	Arg	Val	Gly	
145					150					155					160	
gag	agt	tta	cct	ccg	aga	aag	tct	cat	aga	cgc	tcc	aac	agt	gat	atc	528
Glu	Ser	Leu	Pro	Pro	Arg	Lys	Ser	His	Arg	Arg	Ser	Asn	Ser	Asp	Ile	
				165					170						175	
ccc	agt	ggg	ttt	aat	tcg	atg	cct	ttg	atc	cct	ccg	aga	cca	ttg	gag	576
Pro	Ser	Gly	Phe	Asn	Ser	Met	Pro	Leu	Ile	Pro	Pro	Arg	Pro	Leu	Glu	
			180						185					190		
agg	tct	ttt	tct	ggc	ggg	gag	tgt	gct	gat	tgg	tca	aag	tct	aat	cct	624
Arg	Ser	Phe	Ser	Gly	Gly	Glu	Cys	Ala	Asp	Trp	Ser	Lys	Ser	Asn	Pro	
		195					200					205				
ttt	gtg	aag	aag	gaa	tcg	agc	tgc	gaa	agg	gaa	ggc	gtc	gga	gag	aga	672
Phe	Val	Lys	Lys	Glu	Ser	Ser	Cys	Glu	Arg	Glu	Gly	Val	Gly	Glu	Arg	
	210					215					220					
gaa	gct	atg	gat	gat	ctc	ttc	tca	gca	tat	atg	aat	ctt	gaa	aac	att	720
Glu	Ala	Met	Asp	Asp	Leu	Phe	Ser	Ala	Tyr	Met	Asn	Leu	Glu	Asn	Ile	
225					230					235					240	
gat	gtg	ttg	aac	tcc	tct	gaa	gct	gat	gat	agc	aag	aac	ggc	aat	gag	768
Asp	Val	Leu	Asn	Ser	Ser	Glu	Ala	Asp	Asp	Ser	Lys	Asn	Gly	Asn	Glu	
				245					250						255	
aat	agg	gat	gat	atg	gag	agc	agc	aga	gca	agc	ggg	acc	aag	act	aac	816
Asn	Arg	Asp	Asp	Met	Glu	Ser	Ser	Arg	Ala	Ser	Gly	Thr	Lys	Thr	Asn	
			260					265					270			
ggc	agt	gat	acg	gaa	gga	gag	agc	agc	agt	gtc	aat	gag	agt	gcc	aat	864
Gly	Ser	Asp	Thr	Glu	Gly	Glu	Ser	Ser	Ser	Val	Asn	Glu	Ser	Ala	Asn	
		275					280					285				
aat	aat	atg	aat	tct	tct	ggc	gaa	aag	aga	gag	agc	gtg	aag	aga	aga	912
Asn	Asn	Met	Asn	Ser	Ser	Gly	Glu	Lys	Arg	Glu	Ser	Val	Lys	Arg	Arg	
		290				295					300					
gcg	gct	gga	gga	gat	att	gct	cct	acc	acc	aga	cat	tac	agg	agt	gtt	960
Ala	Ala	Gly	Gly	Asp	Ile	Ala	Pro	Thr	Thr	Arg	His	Tyr	Arg	Ser	Val	
305					310					315					320	
tca	gtg	gac	agt	tgt	ttc	atg	gag	aag	ttg	tct	ttt	ggc	gat	gaa	tct	1008
Ser	Val	Asp	Ser	Cys	Phe	Met	Glu	Lys	Leu	Ser	Phe	Gly	Asp	Glu	Ser	
				325					330					335		
cta	aag	ccg	cct	cct	tct	cct	gga	tct	atg	tca	agg	aaa	gtt	tcc	cct	1056
Leu	Lys	Pro	Pro	Pro	Ser	Pro	Gly	Ser	Met	Ser	Arg	Lys	Val	Ser	Pro	
			340					345					350			
acc</																

```

aat aac ggt gag ttt act gca gcg gaa atg aag aag atc atg gca aat 1152
Asn Asn Gly Glu Phe Thr Ala Ala Glu Met Lys Lys Ile Met Ala Asn
370 375 380

gat aaa cta gca gag atg gcc atg tct gac cct aaa cgt gtc aaa agg 1200
Asp Lys Leu Ala Glu Met Ala Met Ser Asp Pro Lys Arg Val Lys Arg
385 390 395 400

aat gat cct ctt ttc aga atc tta gcg aac cgt caa tcc gca gca cgg 1248
Asn Asp Pro Leu Phe Arg Ile Leu Ala Asn Arg Gln Ser Ala Ala Arg
405 410 415

tca aag gag agg aag atg cgg tac ata gta gaa ttg gaa cac aaa gtg 1296
Ser Lys Glu Arg Lys Met Arg Tyr Ile Val Glu Leu Glu His Lys Val
420 425 430

cag act ctt cag acc gag gct acc aca ttg tct gct cag ctc aca ctt 1344
Gln Thr Leu Gln Thr Glu Ala Thr Thr Leu Ser Ala Gln Leu Thr Leu
435 440 445

ttg cag cgc gat atg atg ggg ttg aca aat cag aac aat gag ctt aag 1392
Leu Gln Arg Asp Met Met Gly Leu Thr Asn Gln Asn Asn Glu Leu Lys
450 455 460

ttc cgg ctt caa gca atg gag caa caa gcg cgt ctt cgc gat gct ctg 1440
Phe Arg Leu Gln Ala Met Glu Gln Gln Ala Arg Leu Arg Asp Ala Leu
465 470 475 480

aac gaa gca ctg aat gga gaa gtc cag cga ctg aaa ctg gca atc ggt 1488
Asn Glu Ala Leu Asn Gly Glu Val Gln Arg Leu Lys Leu Ala Ile Gly
485 490 495

gag agc agt cag aac gaa tca gag aga tca aag atg caa tca ctc aac 1536
Glu Ser Ser Gln Asn Glu Ser Glu Arg Ser Lys Met Gln Ser Leu Asn
500 505 510

gct gag atg ttc cag caa ctc aac atc agc cag tta aga cag cag cca 1584
Ala Glu Met Phe Gln Gln Leu Asn Ile Ser Gln Leu Arg Gln Gln Pro
515 520 525

caa cag atg cag caa cag tct cat cag cag aac cac cag aat gga acc 1632
Gln Gln Met Gln Gln Gln Ser His Gln Gln Asn His Gln Asn Gly Thr
530 535 540

atg gca aca aaa tct gaa tca aat gaa tag 1662
Met Ala Thr Lys Ser Glu Ser Asn Glu *
545 550

```

&lt;210&gt; 116

&lt;211&gt; 553

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (249)...(297)

&lt;223&gt; Conserved domain

&lt;400&gt; 116

```

Met Gly Asp Thr Glu Lys Cys Asn Ser Asp Met Ile Gln Arg Leu His
 1          5          10          15
Ser Ser Phe Gly Thr Thr Ser Ser Ser Ile Pro Lys Asn Pro Ile Ser
 20          25          30
Gln Leu Asp Leu Asn Pro Asn Phe Ile Arg Ser Ser Ala Pro Gln Phe
 35          40          45
Ser Lys Pro Phe Ser Asp Ser Gly Lys Arg Ile Gly Val Pro Pro Ser
 50          55          60
His Pro Asn Leu Ile Pro Pro Thr Ser Pro Phe Ser Gln Ile Pro Thr
 65          70          75          80
Thr Arg Gln Pro Gly Ser His Asn Phe Asn Pro Gly Gly Ala Asn His
 85          90          95
Ser Arg Ser Met Ser Gln Pro Asn Ser Phe Phe Ser Phe Asp Ser Leu
100          105          110
Pro Pro Leu Ser Pro Ser Pro Phe Arg Asp His Asp Val Ser Met Glu
115          120          125
Asp Arg Asp Ser Gly Val Phe Asn Ser Asn His Ser Leu Pro Pro Ser
130          135          140
Pro Phe Thr Arg Cys Asn Ser Thr Ser Ser Ser Ser Leu Arg Val Gly
145          150          155          160
Glu Ser Leu Pro Pro Arg Lys Ser His Arg Arg Ser Asn Ser Asp Ile
165          170          175
Pro Ser Gly Phe Asn Ser Met Pro Leu Ile Pro Pro Arg Pro Leu Glu
180          185          190
Arg Ser Phe Ser Gly Gly Glu Cys Ala Asp Trp Ser Lys Ser Asn Pro
195          200          205
Phe Val Lys Lys Glu Ser Ser Cys Glu Arg Glu Gly Val Gly Glu Arg
210          215          220
Glu Ala Met Asp Asp Leu Phe Ser Ala Tyr Met Asn Leu Glu Asn Ile
225          230          235          240
Asp Val Leu Asn Ser Ser Glu Ala Asp Asp Ser Lys Asn Gly Asn Glu
245          250          255
Asn Arg Asp Asp Met Glu Ser Ser Arg Ala Ser Gly Thr Lys Thr Asn
260          265          270
Gly Ser Asp Thr Glu Gly Glu Ser Ser Ser Val Asn Glu Ser Ala Asn
275          280          285
Asn Asn Met Asn Ser Ser Gly Glu Lys Arg Glu Ser Val Lys Arg Arg
290          295          300
Ala Ala Gly Gly Asp Ile Ala Pro Thr Thr Arg His Tyr Arg Ser Val
305          310          315          320
Ser Val Asp Ser Cys Phe Met Glu Lys Leu Ser Phe Gly Asp Glu Ser
325          330          335
Leu Lys Pro Pro Pro Ser Pro Gly Ser Met Ser Arg Lys Val Ser Pro
340          345          350
Thr Asn Ser Val Asp Gly Asn Ser Gly Ala Ala Phe Ser Ile Glu Phe
355          360          365
Asn Asn Gly Glu Phe Thr Ala Ala Glu Met Lys Lys Ile Met Ala Asn
370          375          380
Asp Lys Leu Ala Glu Met Ala Met Ser Asp Pro Lys Arg Val Lys Arg
385          390          395          400
Asn Asp Pro Leu Phe Arg Ile Leu Ala Asn Arg Gln Ser Ala Ala Arg
405          410          415
Ser Lys Glu Arg Lys Met Arg Tyr Ile Val Glu Leu Glu His Lys Val
420          425          430

```

Gln Thr Leu Gln Thr Glu Ala Thr Thr Leu Ser Ala Gln Leu Thr Leu  
 435 440 445  
 Leu Gln Arg Asp Met Met Gly Leu Thr Asn Gln Asn Asn Glu Leu Lys  
 450 455 460  
 Phe Arg Leu Gln Ala Met Glu Gln Gln Ala Arg Leu Arg Asp Ala Leu  
 465 470 475 480  
 Asn Glu Ala Leu Asn Gly Glu Val Gln Arg Leu Lys Leu Ala Ile Gly  
 485 490 495  
 Glu Ser Ser Gln Asn Glu Ser Glu Arg Ser Lys Met Gln Ser Leu Asn  
 500 505 510  
 Ala Glu Met Phe Gln Gln Leu Asn Ile Ser Gln Leu Arg Gln Gln Pro  
 515 520 525  
 Gln Gln Met Gln Gln Gln Ser His Gln Gln Asn His Gln Asn Gly Thr  
 530 535 540  
 Met Ala Thr Lys Ser Glu Ser Asn Glu  
 545 550

&lt;210&gt; 117

&lt;211&gt; 1218

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1218)

&lt;400&gt; 117

atg gat tct tgg agc tac ggg aga agc gtt ttc atg tca aat gaa act 48  
 Met Asp Ser Trp Ser Tyr Gly Arg Ser Val Phe Met Ser Asn Glu Thr  
 1 5 10 15  
 ctt tta cct tgt gat act ttt gct aag aat aga aga ttt gaa cag aga 96  
 Leu Leu Pro Cys Asp Thr Phe Ala Lys Asn Arg Arg Phe Glu Gln Arg  
 20 25 30  
 ctc tct aat aat gat gat gtg ttg att tct gac atg gct ggt aac tcc 144  
 Leu Ser Asn Asn Asp Asp Val Leu Ile Ser Asp Met Ala Gly Asn Ser  
 35 40 45  
 aat gga ttc agt gct gtt tct att act aaa gtt gtt cct gaa gag gaa 192  
 Asn Gly Phe Ser Ala Val Ser Ile Thr Lys Val Val Pro Glu Glu Glu  
 50 55 60  
 gat gag gag aac ata tct tct tct tca aag ttc tct agt cag gaa ttg 240  
 Asp Glu Glu Asn Ile Ser Ser Ser Ser Lys Phe Ser Ser Gln Glu Leu  
 65 70 75 80  
 aat agg ata gat ttc aaa ctt agg agc ttt ttg gat tta gga aat gat 288  
 Asn Arg Ile Asp Phe Lys Leu Arg Ser Phe Leu Asp Leu Gly Asn Asp  
 85 90 95  
 gat gat gat aca tcc tct aga ggt ttt gct ctg cca tct aaa aag tct 336  
 Asp Asp Asp Thr Ser Ser Arg Gly Phe Ala Leu Pro Ser Lys Lys Ser  
 100 105 110  
 cga gct tca aac ttg tgc tct cag aat ccc ttg tgt caa gtt tat ggg 384



Arg Ala Ser Asn Leu Cys Ser Gln Asn Pro Leu Cys Gln Val Tyr Gly	
115 120 125	
tgt agt aag gat ctg agc tct tcg aaa gat tac cac aaa agg cat aga	432
Cys Ser Lys Asp Leu Ser Ser Ser Lys Asp Tyr His Lys Arg His Arg	
130 135 140	
gtt tgc gag gct cat tcg aaa act tct gtg gtc ata gtt aat ggt ctt	480
Val Cys Glu Ala His Ser Lys Thr Ser Val Val Ile Val Asn Gly Leu	
145 150 155 160	
gaa cag agg ttt tgt caa cag tgc agc agg ttt cat ttc ctc tca gag	528
Glu Gln Arg Phe Cys Gln Gln Cys Ser Arg Phe His Phe Leu Ser Glu	
165 170 175	
ttt gat gat ggc aaa aga agt tgc aga agg cga tta gcc ggt cac aat	576
Phe Asp Asp Gly Lys Arg Ser Cys Arg Arg Arg Leu Ala Gly His Asn	
180 185 190	
gaa cga aga agg aaa cct gca ttc tat ttc cta ccg ggt aag cgc cat	624
Glu Arg Arg Arg Lys Pro Ala Phe Tyr Phe Leu Pro Gly Lys Arg His	
195 200 205	
aag ctt ctt cgc acc tct caa gat gta gta ggc aac aag ttt ctg gag	672
Lys Leu Leu Arg Thr Ser Gln Asp Val Val Gly Asn Lys Phe Leu Glu	
210 215 220	
aat tca tca ttg gta ttg cca gag tca ttt cct ggt agt ctc tta tac	720
Asn Ser Ser Leu Val Leu Pro Glu Ser Phe Pro Gly Ser Leu Leu Tyr	
225 230 235 240	
aga gta ata gat gaa gac gac cac cgt aca agt aga ctc gtg agt ttc	768
Arg Val Ile Asp Glu Asp Asp His Arg Thr Ser Arg Leu Val Ser Phe	
245 250 255	
aaa gat gaa cct act tgt tcc atg ttt cct act aat gag caa aac agc	816
Lys Asp Glu Pro Thr Cys Ser Met Phe Pro Thr Asn Glu Gln Asn Ser	
260 265 270	
agc aga act tat gaa tct aaa cca gca att tat tcc acg gaa gta tcc	864
Ser Arg Thr Tyr Glu Ser Lys Pro Ala Ile Tyr Ser Thr Glu Val Ser	
275 280 285	
tcc att tgg gac tta cat gag acg gcg gca tca cgc tct act cgt gct	912
Ser Ile Trp Asp Leu His Glu Thr Ala Ala Ser Arg Ser Thr Arg Ala	
290 295 300	
ctc tct ctt ctg tca gct cag tcc caa caa cac ttg tct aag ttt cca	960
Leu Ser Leu Leu Ser Ala Gln Ser Gln Gln His Leu Ser Lys Phe Pro	
305 310 315 320	
aac aca acg ttc tca atc acc caa ccc aac caa aat ctc aat cac tca	1008
Asn Thr Thr Phe Ser Ile Thr Gln Pro Asn Gln Asn Leu Asn His Ser	
325 330 335	
tca tca act gac tat cat cag atg gaa caa ccg ttg tgg atc gat cct	1056
Ser Ser Thr Asp Tyr His Gln Met Glu Gln Pro Leu Trp Ile Asp Pro	

```

          340          345          350
ggc aag acc aat tct gct ggt tct agt tct tgt aaa gga aaa gga aca 1104
Gly Lys Thr Asn Ser Ala Gly Ser Ser Ser Cys Lys Gly Lys Gly Thr
          355          360          365

tcc acg gtt gat cta ctg caa ctg tca tca cat ctt caa aga atc gag 1152
Ser Thr Val Asp Leu Leu Gln Leu Ser Ser His Leu Gln Arg Ile Glu
          370          375          380

caa cag agg aat tac act ggt gat gtg aag cag gaa tat aat gag ctt 1200
Gln Gln Arg Asn Tyr Thr Gly Asp Val Lys Gln Glu Tyr Asn Glu Leu
          385          390          395          400

tat ttc cct ggc tcc taa 1218
Tyr Phe Pro Gly Ser *
          405

```

<210> 118  
 <211> 405  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (108)...(233)  
 <223> Conserved domain

```

<400> 118
Met Asp Ser Trp Ser Tyr Gly Arg Ser Val Phe Met Ser Asn Glu Thr
  1          5          10          15
Leu Leu Pro Cys Asp Thr Phe Ala Lys Asn Arg Arg Phe Glu Gln Arg
          20          25          30
Leu Ser Asn Asn Asp Asp Val Leu Ile Ser Asp Met Ala Gly Asn Ser
          35          40          45
Asn Gly Phe Ser Ala Val Ser Ile Thr Lys Val Val Pro Glu Glu Glu
          50          55          60
Asp Glu Glu Asn Ile Ser Ser Ser Lys Phe Ser Ser Gln Glu Leu
          65          70          75          80
Asn Arg Ile Asp Phe Lys Leu Arg Ser Phe Leu Asp Leu Gly Asn Asp
          85          90          95
Asp Asp Asp Thr Ser Ser Arg Gly Phe Ala Leu Pro Ser Lys Lys Ser
          100          105          110
Arg Ala Ser Asn Leu Cys Ser Gln Asn Pro Leu Cys Gln Val Tyr Gly
          115          120          125
Cys Ser Lys Asp Leu Ser Ser Lys Asp Tyr His Lys Arg His Arg
          130          135          140
Val Cys Glu Ala His Ser Lys Thr Ser Val Val Ile Val Asn Gly Leu
          145          150          155          160
Glu Gln Arg Phe Cys Gln Gln Cys Ser Arg Phe His Phe Leu Ser Glu
          165          170          175
Phe Asp Asp Gly Lys Arg Ser Cys Arg Arg Arg Leu Ala Gly His Asn
          180          185          190
Glu Arg Arg Arg Lys Pro Ala Phe Tyr Phe Leu Pro Gly Lys Arg His
          195          200          205
Lys Leu Leu Arg Thr Ser Gln Asp Val Val Gly Asn Lys Phe Leu Glu

```

```

      210              215              220
Asn Ser Ser Leu Val Leu Pro Glu Ser Phe Pro Gly Ser Leu Leu Tyr
225              230              235              240
Arg Val Ile Asp Glu Asp Asp His Arg Thr Ser Arg Leu Val Ser Phe
      245              250              255
Lys Asp Glu Pro Thr Cys Ser Met Phe Pro Thr Asn Glu Gln Asn Ser
      260              265              270
Ser Arg Thr Tyr Glu Ser Lys Pro Ala Ile Tyr Ser Thr Glu Val Ser
      275              280              285
Ser Ile Trp Asp Leu His Glu Thr Ala Ala Ser Arg Ser Thr Arg Ala
      290              295              300
Leu Ser Leu Leu Ser Ala Gln Ser Gln Gln His Leu Ser Lys Phe Pro
305              310              315              320
Asn Thr Thr Phe Ser Ile Thr Gln Pro Asn Gln Asn Leu Asn His Ser
      325              330              335
Ser Ser Thr Asp Tyr His Gln Met Glu Gln Pro Leu Trp Ile Asp Pro
      340              345              350
Gly Lys Thr Asn Ser Ala Gly Ser Ser Ser Cys Lys Gly Lys Gly Thr
      355              360              365
Ser Thr Val Asp Leu Leu Gln Leu Ser Ser His Leu Gln Arg Ile Glu
      370              375              380
Gln Gln Arg Asn Tyr Thr Gly Asp Val Lys Gln Glu Tyr Asn Glu Leu
385              390              395              400
Tyr Phe Pro Gly Ser
      405

```

<210> 119  
 <211> 1913  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (259)...(1593)

```

<400> 119
ggcaagaaga accttagcct ctctttcttc tttctctctc tctctctctg tggtaactgtt 60
ctgtttcaac ttactccct cagtttcaga acaattccct atctagaaga gagataaaac 120
cgagaagggtt ttggagatag aatcttttgt tcttcttttg tccctccttg ctcgattttt 180
gttacgtgtg aagcaataaa aaaaaactga tatagctaaa tcttccatcc attcagagge 240
ttctaaatct gatctgac atg gaa caa gtg ttt gct gat tgg aat ttt gaa 291
              Met Glu Gln Val Phe Ala Asp Trp Asn Phe Glu
              1              5              10

gat aat ttt cac atg tcc act aat aaa aga tca atc aga cca gaa gat 339
Asp Asn Phe His Met Ser Thr Asn Lys Arg Ser Ile Arg Pro Glu Asp
              15              20              25

gaa tta gtg gag.cta ttg tgg aga gat ggt caa gtg gtt tta caa agc 387
Glu Leu Val Glu Leu Leu Trp Arg Asp Gly Gln Val Val Leu Gln Ser
              30              35              40

caa gct cgt aga gaa ccg tca gtc caa gtc caa acc cac aaa caa gaa 435
Gln Ala Arg Arg Glu Pro Ser Val Gln Val Gln Thr His Lys Gln Glu
              45              50              55

```

acc cta aga aaa ccc aac aat att ttt ctt gac aac caa gaa aca gta	483
Thr Leu Arg Lys Pro Asn Asn Ile Phe Leu Asp Asn Gln Glu Thr Val	
60 65 70 75	
caa aag cct aac tac gct gct cta gat gat caa gaa acc gtc tcc tgg	531
Gln Lys Pro Asn Tyr Ala Ala Leu Asp Asp Gln Glu Thr Val Ser Trp	
80 85 90	
ata caa tac cct ccg gat gac gtc atc gac cct ttc gaa tcc gag ttc	579
Ile Gln Tyr Pro Pro Asp Asp Val Ile Asp Pro Phe Glu Ser Glu Phe	
95 100 105	
tcc tct cat ttc ttc tct tcg atc gat cac ctc gga ggt cct gag aag	627
Ser Ser His Phe Phe Ser Ser Ile Asp His Leu Gly Gly Pro Glu Lys	
110 115 120	
cca cga acg atc gaa gag aca gtt aag cat gag gct caa gcc atg gct	675
Pro Arg Thr Ile Glu Glu Thr Val Lys His Glu Ala Gln Ala Met Ala	
125 130 135	
cct cct aag ttt aga tcc tcg gtt ata aca gtc gga ccg agt cat tgc	723
Pro Pro Lys Phe Arg Ser Ser Val Ile Thr Val Gly Pro Ser His Cys	
140 145 150 155	
ggc agc aac cag tca aca aat att cat cag gcc act aca ctt ccg gtt	771
Gly Ser Asn Gln Ser Thr Asn Ile His Gln Ala Thr Thr Leu Pro Val	
160 165 170	
tct atg agt gat aga agc aag aac gtc gaa gaa aga ctt gac act tcg	819
Ser Met Ser Asp Arg Ser Lys Asn Val Glu Glu Arg Leu Asp Thr Ser	
175 180 185	
tca ggt ggc tcc tcc ggt tgc agc tat gga agg aac aac aaa gaa acc	867
Ser Gly Gly Ser Ser Gly Cys Ser Tyr Gly Arg Asn Asn Lys Glu Thr	
190 195 200	
gtt agt gga aca agt gta acc att gac cgt aaa aga aaa cat gtt atg	915
Val Ser Gly Thr Ser Val Thr Ile Asp Arg Lys Arg Lys His Val Met	
205 210 215	
gat gct gat caa gaa tct gtg tct caa tca gat ata ggt ttg acc tca	963
Asp Ala Asp Gln Glu Ser Val Ser Gln Ser Asp Ile Gly Leu Thr Ser	
220 225 230 235	
acc gat gat caa acc atg ggt aac aaa tcg agc caa cgg tca gga tct	1011
Thr Asp Asp Gln Thr Met Gly Asn Lys Ser Ser Gln Arg Ser Gly Ser	
240 245 250	
act cga aga agc cgt gca gct gaa gtt cat aat ctc tca gaa agg agg	1059
Thr Arg Arg Ser Arg Ala Ala Glu Val His Asn Leu Ser Glu Arg Arg	
255 260 265	
agg aga gat cgg atc aat gaa aga atg aaa gct ctt caa gaa ctc ata	1107
Arg Arg Asp Arg Ile Asn Glu Arg Met Lys Ala Leu Gln Glu Leu Ile	
270 275 280	
cct cac tgc agc aga aca gat aaa gct tcg ata ttg gat gaa gca att	1155

Pro His Cys Ser Arg Thr Asp Lys Ala Ser Ile Leu Asp Glu Ala Ile  
 285 290 295

gat tac tta aaa tca ctt caa atg caa ctc caa gtg atg tgg atg gga 1203  
 Asp Tyr Leu Lys Ser Leu Gln Met Gln Leu Val Met Trp Met Gly  
 300 305 310 315

agt gga atg gcg gcg gcg gca gca gca gca gca agt ccg atg atg ttt 1251  
 Ser Gly Met Ala Ala Ala Ala Ala Ala Ala Ser Pro Met Met Phe  
 320 325 330

ccc ggg gta caa tca tct cca tac att aat cag atg gct atg caa agt 1299  
 Pro Gly Val Gln Ser Ser Pro Tyr Ile Asn Gln Met Ala Met Gln Ser  
 335 340 345

cag atg caa ttg tct caa ttc ccg gtt atg aac cgg tcc gct ccg cag 1347  
 Gln Met Gln Leu Ser Gln Phe Pro Val Met Asn Arg Ser Ala Pro Gln  
 350 355 360

aac cat ccc ggt tta gta tgt caa aac ccg gta cag ttg cag ctc caa 1395  
 Asn His Pro Gly Leu Val Cys Gln Asn Pro Val Gln Leu Gln Leu Gln  
 365 370 375

gca cag aac caa atc tta tcg gag cag ctc gct agg tac atg ggc ggg 1443  
 Ala Gln Asn Gln Ile Leu Ser Glu Gln Leu Ala Arg Tyr Met Gly Gly  
 380 385 390 395

att ccc cag atg ccg ccg gcg gga aat cag atg cag acc gtg caa caa 1491  
 Ile Pro Gln Met Pro Pro Ala Gly Asn Gln Met Gln Thr Val Gln Gln  
 400 405 410

caa cca gcg gac atg ttg gga ttt gga tct ccg gcg gga ccg caa agt 1539  
 Gln Pro Ala Asp Met Leu Gly Phe Gly Ser Pro Ala Gly Pro Gln Ser  
 415 420 425

caa ctg tcg gca ccg gcg acc acc gac agt ctt cat atg ggt aaa ata 1587  
 Gln Leu Ser Ala Pro Ala Thr Thr Asp Ser Leu His Met Gly Lys Ile  
 430 435 440

ggc tga cttggcatat agttttcctc cgaaattatt cttcttacag ttggtgattg 1643  
 Gly \*

ttattttattt ttggtgcct aagcaagcat aaaagctaag tcaaattgat tatagagatc 1703  
 taataagtta gtctcactat tataacttat ttttaaacag ttgaattata gtatcaatca 1763  
 agtggtggga acctaaagat catacatgtg tcaatacttt tatatttggt ctcaagggtc 1823  
 atcagaaaaa caaaataaaa aggatagact aggcctgcat ttgacattat catgggcttt 1883  
 tttgggtcta tgaatatgaa cattaacccc 1913

<210> 120

<211> 444

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (253)...(313)

&lt;223&gt; Conserved domain

&lt;400&gt; 120

```

Met Glu Gln Val Phe Ala Asp Trp Asn Phe Glu Asp Asn Phe His Met
 1          5          10          15
Ser Thr Asn Lys Arg Ser Ile Arg Pro Glu Asp Glu Leu Val Glu Leu
 20          25          30
Leu Trp Arg Asp Gly Gln Val Val Leu Gln Ser Gln Ala Arg Arg Glu
 35          40          45
Pro Ser Val Gln Val Gln Thr His Lys Gln Glu Thr Leu Arg Lys Pro
 50          55          60
Asn Asn Ile Phe Leu Asp Asn Gln Glu Thr Val Gln Lys Pro Asn Tyr
 65          70          75          80
Ala Ala Leu Asp Asp Gln Glu Thr Val Ser Trp Ile Gln Tyr Pro Pro
 85          90          95
Asp Asp Val Ile Asp Pro Phe Glu Ser Glu Phe Ser Ser His Phe Phe
100          105          110
Ser Ser Ile Asp His Leu Gly Gly Pro Glu Lys Pro Arg Thr Ile Glu
115          120          125
Glu Thr Val Lys His Glu Ala Gln Ala Met Ala Pro Pro Lys Phe Arg
130          135          140
Ser Ser Val Ile Thr Val Gly Pro Ser His Cys Gly Ser Asn Gln Ser
145          150          155          160
Thr Asn Ile His Gln Ala Thr Thr Leu Pro Val Ser Met Ser Asp Arg
165          170          175
Ser Lys Asn Val Glu Glu Arg Leu Asp Thr Ser Ser Gly Gly Ser Ser
180          185          190
Gly Cys Ser Tyr Gly Arg Asn Asn Lys Glu Thr Val Ser Gly Thr Ser
195          200          205
Val Thr Ile Asp Arg Lys Arg Lys His Val Met Asp Ala Asp Gln Glu
210          215          220
Ser Val Ser Gln Ser Asp Ile Gly Leu Thr Ser Thr Asp Asp Gln Thr
225          230          235          240
Met Gly Asn Lys Ser Ser Gln Arg Ser Gly Ser Thr Arg Arg Ser Arg
245          250          255
Ala Ala Glu Val His Asn Leu Ser Glu Arg Arg Arg Arg Asp Arg Ile
260          265          270
Asn Glu Arg Met Lys Ala Leu Gln Glu Leu Ile Pro His Cys Ser Arg
275          280          285
Thr Asp Lys Ala Ser Ile Leu Asp Glu Ala Ile Asp Tyr Leu Lys Ser
290          295          300
Leu Gln Met Gln Leu Gln Val Met Trp Met Gly Ser Gly Met Ala Ala
305          310          315          320
Ala Ala Ala Ala Ala Ala Ser Pro Met Met Phe Pro Gly Val Gln Ser
325          330          335
Ser Pro Tyr Ile Asn Gln Met Ala Met Gln Ser Gln Met Gln Leu Ser
340          345          350
Gln Phe Pro Val Met Asn Arg Ser Ala Pro Gln Asn His Pro Gly Leu
355          360          365
Val Cys Gln Asn Pro Val Gln Leu Gln Leu Gln Ala Gln Asn Gln Ile
370          375          380
Leu Ser Glu Gln Leu Ala Arg Tyr Met Gly Gly Ile Pro Gln Met Pro
385          390          395          400
Pro Ala Gly Asn Gln Met Gln Thr Val Gln Gln Gln Pro Ala Asp Met
405          410          415
Leu Gly Phe Gly Ser Pro Ala Gly Pro Gln Ser Gln Leu Ser Ala Pro
420          425          430

```



cct gga gct tgg acc tcg gtg tct ggt gag gag gaa gaa gcg cgt ggg 583  
 Pro Gly Ala Trp Thr Ser Val Ser Gly Glu Glu Glu Glu Ala Arg Gly  
 165 170 175

aaa aaa gga ggt gac gac gga aac taa taagtgaat acgttttagg 630  
 Lys Lys Gly Gly Asp Asp Gly Asn \*  
 180 185

gtattttcaa gggaatatgt agtaaatagt catggatc 668

<210> 122  
 <211> 187  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (30)...(164)  
 <223> Conserved domain

<400> 122  
 Met Glu Asn Asn Asn Gly Asn Asn Gln Leu Pro Pro Lys Gly Asn Glu  
 1 5 10 15  
 Gln Leu Lys Ser Phe Trp Ser Lys Glu Met Glu Gly Asn Leu Asp Phe  
 20 25 30  
 Lys Asn His Asp Leu Pro Ile Thr Arg Ile Lys Lys Ile Met Lys Tyr  
 35 40 45  
 Asp Pro Asp Val Thr Met Ile Ala Ser Glu Ala Pro Ile Leu Leu Ser  
 50 55 60  
 Lys Ala Cys Glu Met Phe Ile Met Asp Leu Thr Met Arg Ser Trp Leu  
 65 70 75 80  
 His Ala Gln Glu Ser Lys Arg Val Thr Leu Gln Lys Ser Asn Val Asp  
 85 90 95  
 Ala Ala Val Ala Gln Thr Val Ile Phe Asp Phe Leu Leu Asp Asp  
 100 105 110  
 Ile Glu Val Lys Arg Glu Ser Val Ala Ala Ala Asp Pro Val Ala  
 115 120 125  
 Met Pro Pro Ile Asp Asp Gly Glu Leu Pro Pro Gly Met Val Ile Gly  
 130 135 140  
 Thr Pro Val Cys Cys Ser Leu Gly Ile His Gln Pro Gln Pro Gln Met  
 145 150 155 160  
 Gln Ala Trp Pro Gly Ala Trp Thr Ser Val Ser Gly Glu Glu Glu Glu  
 165 170 175  
 Ala Arg Gly Lys Lys Gly Gly Asp Asp Gly Asn  
 180 185

<210> 123  
 <211> 1501  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (111)...(1421)



&lt;400&gt; 123

```

tcgacatcag atttctctca cggattccta atcattttta ttatatattgg atatttgcta 60
attcttcccg tgtataaatc tcatataaac acgcatacata catatatatt atg tgc 116
                                         Met Cys
                                         1

agc gtc ttt gag ttt caa gac atg gac aac ttc caa gga gat cta aca 164
Ser Val Phe Glu Phe Gln Asp Met Asp Asn Phe Gln Gly Asp Leu Thr
      5              10              15

gac gtc gta cga gga ata gga tca ggc cac gtg tca cca tct cct gga 212
Asp Val Val Arg Gly Ile Gly Ser Gly His Val Ser Pro Ser Pro Gly
      20              25              30

cca ccg gaa ggt cca tct ccg agc agc atg tct ccg ccg cca aca tca 260
Pro Pro Glu Gly Pro Ser Pro Ser Ser Met Ser Pro Pro Pro Thr Ser
      35              40              45              50

gat ctc cac gtg gaa ttc ccc tcc gcc gct act tct gcc agc tgt ctc 308
Asp Leu His Val Glu Phe Pro Ser Ala Ala Thr Ser Ala Ser Cys Leu
      55              60              65

gca aat ccc ttc gga gac ccg ttc gta agc atg aag gat cct ctc atc 356
Ala Asn Pro Phe Gly Asp Pro Phe Val Ser Met Lys Asp Pro Leu Ile
      70              75              80

cac ctc ccg gcc agc tac atc tcc ggc gcc ggt gat aat aaa agc aac 404
His Leu Pro Ala Ser Tyr Ile Ser Gly Ala Gly Asp Asn Lys Ser Asn
      85              90              95

aaa agt ttt gca atc ttt cca aag att ttt gag gat gat cat att aag 452
Lys Ser Phe Ala Ile Phe Pro Lys Ile Phe Glu Asp Asp His Ile Lys
      100              105              110

agt caa tgc agt gtc ttc cca aga att aag atc tcg caa agt aac aat 500
Ser Gln Cys Ser Val Phe Pro Arg Ile Lys Ile Ser Gln Ser Asn Asn
      115              120              125              130

atc cac gat gcc tcc acg tgt aat tct ccg gcc ata acc gtc tcc tct 548
Ile His Asp Ala Ser Thr Cys Asn Ser Pro Ala Ile Thr Val Ser Ser
      135              140              145

gcc gcc gta gca gct tcg ccg tgg ggc atg atc aac gtt aat acc act 596
Ala Ala Val Ala Ala Ser Pro Trp Gly Met Ile Asn Val Asn Thr Thr
      150              155              160

aac agt cca aga aac tgt tta ctt gtc gat aat aat aac aac acg tca 644
Asn Ser Pro Arg Asn Cys Leu Leu Val Asp Asn Asn Asn Asn Thr Ser
      165              170              175

tca tgc tca cag gtt cag atc tct tct tcc cct cgg aat ctc gga att 692
Ser Cys Ser Gln Val Gln Ile Ser Ser Ser Pro Arg Asn Leu Gly Ile
      180              185              190

aag aga agg aag agc cag gca aag aaa gtg gtg tgc ata ccg gct cca 740
Lys Arg Arg Lys Ser Gln Ala Lys Lys Val Val Cys Ile Pro Ala Pro
      195              200              205              210

```

gcc gct atg aac agc cgg tcc agt gga gaa gtt gtt ccg tct gat cta	788
Ala Ala Met Asn Ser Arg Ser Ser Gly Glu Val Val Pro Ser Asp Leu	
215 220 225	
tgg gct tgg cga aag tac ggt caa aaa cct atc aaa ggt tct cct tat	836
Trp Ala Trp Arg Lys Tyr Gly Gln Lys Pro Ile Lys Gly Ser Pro Tyr	
230 235 240	
cca agg ggt tac tac aga tgt agc agc tca aaa ggt tgt tca gct agg	884
Pro Arg Gly Tyr Tyr Arg Cys Ser Ser Ser Lys Gly Cys Ser Ala Arg	
245 250 255	
aaa caa gtc gaa cgt agc cgc act gat cca aac atg tta gtc att act	932
Lys Gln Val Glu Arg Ser Arg Thr Asp Pro Asn Met Leu Val Ile Thr	
260 265 270	
tac acc tct gag cat aac cac cca tgg cct act caa cgc aac gct ctc	980
Tyr Thr Ser Glu His Asn His Pro Trp Pro Thr Gln Arg Asn Ala Leu	
275 280 285 290	
gca ggt tcc act cgt tcc tct tcc tcc tcc tct tta aac cct tct tcc	1028
Ala Gly Ser Thr Arg Ser Ser Ser Ser Ser Ser Leu Asn Pro Ser Ser	
295 300 305	
aaa tcc tca acc gca gcc gcc act act tct ccc tca tcc aga gtt ttc	1076
Lys Ser Ser Thr Ala Ala Ala Thr Thr Ser Pro Ser Ser Arg Val Phe	
310 315 320	
caa aac aac agc agc aaa gac gaa ccc aat aac tcc aac ttg cct tcc	1124
Gln Asn Asn Ser Ser Lys Asp Glu Pro Asn Asn Ser Asn Leu Pro Ser	
325 330 335	
tct tcc act cat cct cct ttt gac gcc gcc gca att aag gag gag aac	1172
Ser Ser Thr His Pro Pro Phe Asp Ala Ala Ala Ile Lys Glu Glu Asn	
340 345 350	
gtg gaa gag cgt cag gaa aag atg gag ttc gat tat aat gac gtt gaa	1220
Val Glu Glu Arg Gln Glu Lys Met Glu Phe Asp Tyr Asn Asp Val Glu	
355 360 365 370	
aat acc tat aga ccg gag ttg ttg caa gag ttt caa cat cag ccg gag	1268
Asn Thr Tyr Arg Pro Glu Leu Leu Gln Glu Phe Gln His Gln Pro Glu	
375 380 385	
gat ttc ttt gcc gat ctc gac gag ctt gag gga gat tct ttg act atg	1316
Asp Phe Phe Ala Asp Leu Asp Glu Leu Glu Gly Asp Ser Leu Thr Met	
390 395 400	
ttg ctc tct cac agt agc ggc gga ggc aac atg gaa aac aaa acg acg	1364
Leu Leu Ser His Ser Ser Gly Gly Gly Asn Met Glu Asn Lys Thr Thr	
405 410 415	
att cca gac gtt ttt agt gat ttc ttt gac gac gac gag tcc tca agg	1412
Ile Pro Asp Val Phe Ser Asp Phe Phe Asp Asp Asp Glu Ser Ser Arg	
420 425 430	

tcg tta taa atattgttgt taatgtatac atagaaatga aattattcat 1461  
 Ser Leu \*  
 435

gtaattcgtt ttgtgttaaa tgacggtatt tgcctttgca 1501

<210> 124  
 <211> 436  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (223)...(283)  
 <223> Conserved domain

<400> 124  
 Met Cys Ser Val Phe Glu Phe Gln Asp Met Asp Asn Phe Gln Gly Asp  
 1 5 10 15  
 Leu Thr Asp Val Val Arg Gly Ile Gly Ser Gly His Val Ser Pro Ser  
 20 25 30  
 Pro Gly Pro Pro Glu Gly Pro Ser Pro Ser Ser Met Ser Pro Pro Pro  
 35 40 45  
 Thr Ser Asp Leu His Val Glu Phe Pro Ser Ala Ala Thr Ser Ala Ser  
 50 55 60  
 Cys Leu Ala Asn Pro Phe Gly Asp Pro Phe Val Ser Met Lys Asp Pro  
 65 70 75 80  
 Leu Ile His Leu Pro Ala Ser Tyr Ile Ser Gly Ala Gly Asp Asn Lys  
 85 90 95  
 Ser Asn Lys Ser Phe Ala Ile Phe Pro Lys Ile Phe Glu Asp Asp His  
 100 105 110  
 Ile Lys Ser Gln Cys Ser Val Phe Pro Arg Ile Lys Ile Ser Gln Ser  
 115 120 125  
 Asn Asn Ile His Asp Ala Ser Thr Cys Asn Ser Pro Ala Ile Thr Val  
 130 135 140  
 Ser Ser Ala Ala Val Ala Ala Ser Pro Trp Gly Met Ile Asn Val Asn  
 145 150 155 160  
 Thr Thr Asn Ser Pro Arg Asn Cys Leu Leu Val Asp Asn Asn Asn Asn  
 165 170 175  
 Thr Ser Ser Cys Ser Gln Val Gln Ile Ser Ser Ser Pro Arg Asn Leu  
 180 185 190  
 Gly Ile Lys Arg Arg Lys Ser Gln Ala Lys Lys Val Val Cys Ile Pro  
 195 200 205  
 Ala Pro Ala Ala Met Asn Ser Arg Ser Ser Gly Glu Val Val Pro Ser  
 210 215 220  
 Asp Leu Trp Ala Trp Arg Lys Tyr Gly Gln Lys Pro Ile Lys Gly Ser  
 225 230 235 240  
 Pro Tyr Pro Arg Gly Tyr Tyr Arg Cys Ser Ser Ser Lys Gly Cys Ser  
 245 250 255  
 Ala Arg Lys Gln Val Glu Arg Ser Arg Thr Asp Pro Asn Met Leu Val  
 260 265 270  
 Ile Thr Tyr Thr Ser Glu His Asn His Pro Trp Pro Thr Gln Arg Asn  
 275 280 285  
 Ala Leu Ala Gly Ser Thr Arg Ser Ser Ser Ser Ser Ser Leu Asn Pro  
 290 295 300  
 Ser Ser Lys Ser Ser Thr Ala Ala Ala Thr Thr Ser Pro Ser Ser Arg  
 305 310 315 320

Val Phe Gln Asn Asn Ser Ser Lys Asp Glu Pro Asn Asn Ser Asn Leu  
 325 330 335  
 Pro Ser Ser Ser Thr His Pro Pro Phe Asp Ala Ala Ala Ile Lys Glu  
 340 345 350  
 Glu Asn Val Glu Glu Arg Gln Glu Lys Met Glu Phe Asp Tyr Asn Asp  
 355 360 365  
 Val Glu Asn Thr Tyr Arg Pro Glu Leu Leu Gln Glu Phe Gln His Gln  
 370 375 380  
 Pro Glu Asp Phe Phe Ala Asp Leu Asp Glu Leu Glu Gly Asp Ser Leu  
 385 390 395 400  
 Thr Met Leu Leu Ser His Ser Ser Gly Gly Gly Asn Met Glu Asn Lys  
 405 410 415  
 Thr Thr Ile Pro Asp Val Phe Ser Asp Phe Phe Asp Asp Asp Glu Ser  
 420 425 430  
 Ser Arg Ser Leu  
 435

&lt;210&gt; 125

&lt;211&gt; 1828

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (36)...(1622)

&lt;400&gt; 125

gagagctaataaaaaatttta tcaaagaaga ctaat atg gag aag gac gat ttc 53  
 Met Glu Lys Asp Asp Phe  
 1 5

ttg agg agt ggt cat gga aga gaa gaa agc cat gat gag atg aga aaa 101  
 Leu Arg Ser Gly His Gly Arg Glu Glu Ser His Asp Glu Met Arg Lys  
 10 15 20

ctt gat tca tct cac gat gat tct cat caa gaa cac gac cat att ata 149  
 Leu Asp Ser Ser His Asp Asp Ser His Gln Glu His Asp His Ile Ile  
 25 30 35

aga tcc aag ttg gac tca act aaa gtc gaa atg gat gag gct aaa gag 197  
 Arg Ser Lys Leu Asp Ser Thr Lys Val Glu Met Asp Glu Ala Lys Glu  
 40 45 50

gaa aat cga aga cta aag tca tca ttg agt aaa atc aag aaa gat ttt 245  
 Glu Asn Arg Arg Leu Lys Ser Ser Leu Ser Lys Ile Lys Lys Asp Phe  
 55 60 65 70

gac atc ctt caa aca caa tac aac caa tta atg gcc aaa cat aac gaa 293  
 Asp Ile Leu Gln Thr Gln Tyr Asn Gln Leu Met Ala Lys His Asn Glu  
 75 80 85

cca acc aag ttc caa tca aaa ggg cat cat caa gac aaa ggc gaa gat 341  
 Pro Thr Lys Phe Gln Ser Lys Gly His His Gln Asp Lys Gly Glu Asp  
 90 95 100

gaa gac aga gaa aaa gtt aac gaa cgt gaa gaa ctt gtc tcg ttg agc 389

Glu Asp Arg Glu Lys Val Asn Glu Arg Glu Glu Leu Val Ser Leu Ser	
105 110 115	
cta ggc aga cgg tta aat tca gag gtt cca agt ggt tcg aat aaa gaa	437
Leu Gly Arg Arg Leu Asn Ser Glu Val Pro Ser Gly Ser Asn Lys Glu	
120 125 130	
gaa aaa aat aaa gat gtt gaa gaa gcg gaa ggt gac aga aat tat gat	485
Glu Lys Asn Lys Asp Val Glu Glu Ala Glu Gly Asp Arg Asn Tyr Asp	
135 140 145 150	
gat aat gaa aaa agc agt att caa ggg ttg agt atg ggg att gaa tac	533
Asp Asn Glu Lys Ser Ser Ile Gln Gly Leu Ser Met Gly Ile Glu Tyr	
155 160 165	
aag gct ttg agt aat cct aat gag aag tta gag att gat cat aat caa	581
Lys Ala Leu Ser Asn Pro Asn Glu Lys Leu Glu Ile Asp His Asn Gln	
170 175 180	
gaa acc atg tcg ttg gag att agt aac aat aat aag atc aga tca caa	629
Glu Thr Met Ser Leu Glu Ile Ser Asn Asn Asn Lys Ile Arg Ser Gln	
185 190 195	
aat agt ttt ggg ttt aag aat gat gga gat gat cat gaa gat gaa gat	677
Asn Ser Phe Gly Phe Lys Asn Asp Gly Asp Asp His Glu Asp Glu Asp	
200 205 210	
gag att ttg cct caa aac ctt gtt aag aaa act agg gtt tcg gtg aga	725
Glu Ile Leu Pro Gln Asn Leu Val Lys Lys Thr Arg Val Ser Val Arg	
215 220 225 230	
tca aga tgt gag aca cca acg atg aac gac gga tgt caa tgg agg aaa	773
Ser Arg Cys Glu Thr Pro Thr Met Asn Asp Gly Cys Gln Trp Arg Lys	
235 240 245	
tat ggc cag aaa ata gct aaa ggc aat cca tgt ccc cga gct tac tat	821
Tyr Gly Gln Lys Ile Ala Lys Gly Asn Pro Cys Pro Arg Ala Tyr Tyr	
250 255 260	
cgt tgc acc att gca gct tct tgt cca gta aga aaa cag gtg caa aga	869
Arg Cys Thr Ile Ala Ala Ser Cys Pro Val Arg Lys Gln Val Gln Arg	
265 270 275	
tgt tca gaa gat atg tct ata ctt atc tca acg tac gaa gga aca cat	917
Cys Ser Glu Asp Met Ser Ile Leu Ile Ser Thr Tyr Glu Gly Thr His	
280 285 290	
aac cat cca ctt ccc atg tca gca act gcc atg gcc tct gcc act tcc	965
Asn His Pro Leu Pro Met Ser Ala Thr Ala Met Ala Ser Ala Thr Ser	
295 300 305 310	
gct gcc gcc tcc atg ctt ctc tcc ggc gcc tcc tcc tcc tca tcc gcc	1013
Ala Ala Ala Ser Met Leu Leu Ser Gly Ala Ser Ser Ser Ser Ala	
315 320 325	
gca gct gat ctt cat ggc ctt aac ttc tct ctt tcc ggc aac aac atc	1061
Ala Ala Asp Leu His Gly Leu Asn Phe Ser Leu Ser Gly Asn Asn Ile	

330	335	340	
act cca aaa cct aaa act cat ttc ctc caa tcc cct tct tct tct ggc			1109
Thr Pro Lys Pro Lys Thr His Phe Leu Gln Ser Pro Ser Ser Ser Gly			
345	350	355	
cat ccg acc gtc act ctc gac ctc aca acc tcc tcc tcg tcg cag caa			1157
His Pro Thr Val Thr Leu Asp Leu Thr Thr Ser Ser Ser Ser Gln Gln			
360	365	370	
ccg ttc tta tca atg ctc aat aga ttc agc tct cct cca agt aat gtc			1205
Pro Phe Leu Ser Met Leu Asn Arg Phe Ser Ser Pro Pro Ser Asn Val			
375	380	385	390
tca cga tct aat agt tat cct tca acc aat ctc aac ttt tca aac aac			1253
Ser Arg Ser Asn Ser Tyr Pro Ser Thr Asn Leu Asn Phe Ser Asn Asn			
395	400	405	
acc aac aca ttg atg aat tgg ggt ggt ggt ggt aat ccc agt gat caa			1301
Thr Asn Thr Leu Met Asn Trp Gly Gly Gly Gly Asn Pro Ser Asp Gln			
410	415	420	
tac cgt gca gct tac ggc aac att aac acc cat cag caa tca cct tac			1349
Tyr Arg Ala Ala Tyr Gly Asn Ile Asn Thr His Gln Gln Ser Pro Tyr			
425	430	435	
cac aaa atc att caa acc cga acc gcc ggg tca tct ttc gat ccg ttt			1397
His Lys Ile Ile Gln Thr Arg Thr Ala Gly Ser Ser Phe Asp Pro Phe			
440	445	450	
gga aga tca tct tca tca cat tct cca caa ata aat ctt gat cat atc			1445
Gly Arg Ser Ser Ser Ser His Ser Pro Gln Ile Asn Leu Asp His Ile			
455	460	465	470
gga atc aag aac atc atc agt cac caa gtg cca tct tta ccg gct gaa			1493
Gly Ile Lys Asn Ile Ile Ser His Gln Val Pro Ser Leu Pro Ala Glu			
475	480	485	
aca atc aag gca atc acg aca gat cca agt ttc caa tcg gct ttg gcg			1541
Thr Ile Lys Ala Ile Thr Thr Asp Pro Ser Phe Gln Ser Ala Leu Ala			
490	495	500	
aca gct cta tct tcc atc atg ggc ggc gat tta aag att gat cac aat			1589
Thr Ala Leu Ser Ser Ile Met Gly Gly Asp Leu Lys Ile Asp His Asn			
505	510	515	
gtg act aga aat gaa gct gag aag agc cct taa agagaattgt tatatatatg			1642
Val Thr Arg Asn Glu Ala Glu Lys Ser Pro *			
520	525		
tttttatata ctcagtagcat tggtaaatgg gtttagactt tcaactagttt cctagttcat			1702
ctatatattg gttgtttaat cacaagttaa ttttggtgtt ggagtttatg gaactaatgt			1762
gtacatatga aacttttagaa cgaataaata aaacttgga ttccttttta aaaaaaaaaa			1822
aaaaaa			1828
<210> 126			
<211> 528			

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (239)...(298)

&lt;223&gt; Conserved domain

&lt;400&gt; 126

```

Met Glu Lys Asp Asp Phe Leu Arg Ser Gly His Gly Arg Glu Glu Ser
 1          5          10          15
His Asp Glu Met Arg Lys Leu Asp Ser Ser His Asp Asp Ser His Gln
          20          25          30
Glu His Asp His Ile Ile Arg Ser Lys Leu Asp Ser Thr Lys Val Glu
          35          40          45
Met Asp Glu Ala Lys Glu Glu Asn Arg Arg Leu Lys Ser Ser Leu Ser
          50          55          60
Lys Ile Lys Lys Asp Phe Asp Ile Leu Gln Thr Gln Tyr Asn Gln Leu
          65          70          75          80
Met Ala Lys His Asn Glu Pro Thr Lys Phe Gln Ser Lys Gly His His
          85          90          95
Gln Asp Lys Gly Glu Asp Glu Asp Arg Glu Lys Val Asn Glu Arg Glu
          100          105          110
Glu Leu Val Ser Leu Ser Leu Gly Arg Arg Leu Asn Ser Glu Val Pro
          115          120          125
Ser Gly Ser Asn Lys Glu Glu Lys Asn Lys Asp Val Glu Glu Ala Glu
          130          135          140
Gly Asp Arg Asn Tyr Asp Asp Asn Glu Lys Ser Ser Ile Gln Gly Leu
          145          150          155          160
Ser Met Gly Ile Glu Tyr Lys Ala Leu Ser Asn Pro Asn Glu Lys Leu
          165          170          175
Glu Ile Asp His Asn Gln Glu Thr Met Ser Leu Glu Ile Ser Asn Asn
          180          185          190
Asn Lys Ile Arg Ser Gln Asn Ser Phe Gly Phe Lys Asn Asp Gly Asp
          195          200          205
Asp His Glu Asp Glu Asp Glu Ile Leu Pro Gln Asn Leu Val Lys Lys
          210          215          220
Thr Arg Val Ser Val Arg Ser Arg Cys Glu Thr Pro Thr Met Asn Asp
          225          230          235          240
Gly Cys Gln Trp Arg Lys Tyr Gly Gln Lys Ile Ala Lys Gly Asn Pro
          245          250          255
Cys Pro Arg Ala Tyr Tyr Arg Cys Thr Ile Ala Ala Ser Cys Pro Val
          260          265          270
Arg Lys Gln Val Gln Arg Cys Ser Glu Asp Met Ser Ile Leu Ile Ser
          275          280          285
Thr Tyr Glu Gly Thr His Asn His Pro Leu Pro Met Ser Ala Thr Ala
          290          295          300
Met Ala Ser Ala Thr Ser Ala Ala Ala Ser Met Leu Leu Ser Gly Ala
          305          310          315          320
Ser Ser Ser Ser Ser Ala Ala Ala Asp Leu His Gly Leu Asn Phe Ser
          325          330          335
Leu Ser Gly Asn Asn Ile Thr Pro Lys Pro Lys Thr His Phe Leu Gln
          340          345          350
Ser Pro Ser Ser Ser Gly His Pro Thr Val Thr Leu Asp Leu Thr Thr
          355          360          365
Ser Ser Ser Ser Gln Gln Pro Phe Leu Ser Met Leu Asn Arg Phe Ser
          370          375          380

```

Ser Pro Pro Ser Asn Val Ser Arg Ser Asn Ser Tyr Pro Ser Thr Asn  
 385 390 395 400  
 Leu Asn Phe Ser Asn Asn Thr Asn Thr Leu Met Asn Trp Gly Gly Gly  
 405 410 415  
 Gly Asn Pro Ser Asp Gln Tyr Arg Ala Ala Tyr Gly Asn Ile Asn Thr  
 420 425 430  
 His Gln Gln Ser Pro Tyr His Lys Ile Ile Gln Thr Arg Thr Ala Gly  
 435 440 445  
 Ser Ser Phe Asp Pro Phe Gly Arg Ser Ser Ser Ser His Ser Pro Gln  
 450 455 460  
 Ile Asn Leu Asp His Ile Gly Ile Lys Asn Ile Ile Ser His Gln Val  
 465 470 475 480  
 Pro Ser Leu Pro Ala Glu Thr Ile Lys Ala Ile Thr Thr Asp Pro Ser  
 485 490 495  
 Phe Gln Ser Ala Leu Ala Thr Ala Leu Ser Ser Ile Met Gly Gly Asp  
 500 505 510  
 Leu Lys Ile Asp His Asn Val Thr Arg Asn Glu Ala Glu Lys Ser Pro  
 515 520 525

&lt;210&gt; 127

&lt;211&gt; 913

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (81)...(761)

&lt;400&gt; 127

agaaaacatc tctcactctc taaaatacac actctcatca aaaaccttct ctccggttca 60  
 gaagcattca agaattccatt atg agc tca tct gat tcc gtt aat aac ggc gtt 113  
 Met Ser Ser Ser Asp Ser Val Asn Asn Gly Val  
 1 5 10

aac tca cgg atg tac ttc cgt aac ccg agt ttc agc aac gtt atc tta 161  
 Asn Ser Arg Met Tyr Phe Arg Asn Pro Ser Phe Ser Asn Val Ile Leu  
 15 20 25

aac gat aac tgg agc gac ttg ccg tta agt gtc gac gat tct caa gac 209  
 Asn Asp Asn Trp Ser Asp Leu Pro Leu Ser Val Asp Asp Ser Gln Asp  
 30 35 40

atg gct att tac aac act ctc cgt gat gcc gtt agc tcc ggc tgg aca 257  
 Met Ala Ile Tyr Asn Thr Leu Arg Asp Ala Val Ser Ser Gly Trp Thr  
 45 50 55

ccc tcc gtt cct ccc gtt acc tct ccg gcg gag gaa aat aag cct ccg 305  
 Pro Ser Val Pro Pro Val Thr Ser Pro Ala Glu Glu Asn Lys Pro Pro  
 60 65 70 75

gcg acg aag gcg agt ggc tca cac gcg ccg agg cag aag ggg atg cag 353  
 Ala Thr Lys Ala Ser Gly Ser His Ala Pro Arg Gln Lys Gly Met Gln  
 80 85 90

tac aga gga gtg agg agg agg ccg tgg ggg aaa ttc gcg gcg gag att 401  
 Tyr Arg Gly Val Arg Arg Arg Pro Trp Gly Lys Phe Ala Ala Glu Ile



```

          95              100              105
agg gat ccg aag aag aac gga gct agg gtt tgg ctc ggg act tac gag 449
Arg Asp Pro Lys Lys Asn Gly Ala Arg Val Trp Leu Gly Thr Tyr Glu
      110              115              120

acg ccg gag gac gcg gcg gtg gcg tac gac cga gcg gcg ttt cag ctc 497
Thr Pro Glu Asp Ala Ala Val Ala Tyr Asp Arg Ala Ala Phe Gln Leu
      125              130              135

aga gga tcg aaa gct aag ctg aat ttt ccg cat ttg att ggt tct tgt 545
Arg Gly Ser Lys Ala Lys Leu Asn Phe Pro His Leu Ile Gly Ser Cys
      140              145              150              155

aag tat gag ccg gtt agg att agg cct cgc cgt cgc tcg ccg gaa ccg 593
Lys Tyr Glu Pro Val Arg Ile Arg Pro Arg Arg Arg Ser Pro Glu Pro
      160              165              170

tca gtc tcc gat cag tta acg tcg gag cag aag agg gaa agc cac gtg 641
Ser Val Ser Asp Gln Leu Thr Ser Glu Gln Lys Arg Glu Ser His Val
      175              180              185

gat gac ggc gag tct agt ttg gtt gta ccg gag ttg gat ttc acg gtg 689
Asp Asp Gly Glu Ser Ser Leu Val Val Pro Glu Leu Asp Phe Thr Val
      190              195              200

gat cag ttt tac ttc gat ggt agt tta tta atg gac caa tca gaa tgt 737
Asp Gln Phe Tyr Phe Asp Gly Ser Leu Leu Met Asp Gln Ser Glu Cys
      205              210              215

tct tat tct gat aat ccg ata taa ttagtttttaa gattaagcaa aatttgtcca 791
Ser Tyr Ser Asp Asn Arg Ile *
      220              225

acgagttttg ctgtatgaaa tatctatcga tgactcaaca ggttttgatc atgatcatat 851
gtaatgtgat ggaaattaaa tattgacggtt tgtttttttg ttgtaaaaaa aaaaaaaaaa 911
aa                                                                    913

```

&lt;210&gt; 128

&lt;211&gt; 226

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (89)...(157)

&lt;223&gt; Conserved domain

&lt;400&gt; 128

```

Met Ser Ser Ser-Asp Ser Val Asn Asn Gly Val Asn Ser Arg Met Tyr
  1              5              10              15
Phe Arg Asn Pro Ser Phe Ser Asn Val Ile Leu Asn Asp Asn Trp Ser
      20              25              30
Asp Leu Pro Leu Ser Val Asp Asp Ser Gln Asp Met Ala Ile Tyr Asn
      35              40              45
Thr Leu Arg Asp Ala Val Ser Ser Gly Trp Thr Pro Ser Val Pro Pro
      50              55              60

```

Val Thr Ser Pro Ala Glu Glu Asn Lys Pro Pro Ala Thr Lys Ala Ser  
 65 70 75 80  
 Gly Ser His Ala Pro Arg Gln Lys Gly Met Gln Tyr Arg Gly Val Arg  
 85 90 95  
 Arg Arg Pro Trp Gly Lys Phe Ala Ala Glu Ile Arg Asp Pro Lys Lys  
 100 105 110  
 Asn Gly Ala Arg Val Trp Leu Gly Thr Tyr Glu Thr Pro Glu Asp Ala  
 115 120 125  
 Ala Val Ala Tyr Asp Arg Ala Ala Phe Gln Leu Arg Gly Ser Lys Ala  
 130 135 140  
 Lys Leu Asn Phe Pro His Leu Ile Gly Ser Cys Lys Tyr Glu Pro Val  
 145 150 155 160  
 Arg Ile Arg Pro Arg Arg Ser Pro Glu Pro Ser Val Ser Asp Gln  
 165 170 175  
 Leu Thr Ser Glu Gln Lys Arg Glu Ser His Val Asp Asp Gly Glu Ser  
 180 185 190  
 Ser Leu Val Val Pro Glu Leu Asp Phe Thr Val Asp Gln Phe Tyr Phe  
 195 200 205  
 Asp Gly Ser Leu Leu Met Asp Gln Ser Glu Cys Ser Tyr Ser Asp Asn  
 210 215 220  
 Arg Ile  
 225

&lt;210&gt; 129

&lt;211&gt; 1338

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1338)

&lt;400&gt; 129

atg tcg ttt cga ggc att gtt caa gat ttg aga gat ggg ttt ggg agc 48  
 Met Ser Phe Arg Gly Ile Val Gln Asp Leu Arg Asp Gly Phe Gly Ser  
 1 5 10 15  
 ttg tca agg agg agt ttc gat ttt agg ctc tcg agt ctt cat aaa ggg 96  
 Leu Ser Arg Arg Ser Phe Asp Phe Arg Leu Ser Ser Leu His Lys Gly  
 20 25 30  
 aaa gct cag ggt tct tcg ttc cgt gag tat tcg tca tcc cgt gat ctc 144  
 Lys Ala Gln Gly Ser Ser Phe Arg Glu Tyr Ser Ser Ser Arg Asp Leu  
 35 40 45  
 ttg tcg cct gtg ata gtt cag aca agt aga tgg gct aat ctt cct cca 192  
 Leu Ser Pro Val Ile Val Gln Thr Ser Arg Trp Ala Asn Leu Pro Pro  
 50 55 60  
 gag tta ctc ttt gat gtg atc aaa aga tta gag gaa agt gag agt aat 240  
 Glu Leu Leu Phe Asp Val Ile Lys Arg Leu Glu Glu Ser Glu Ser Asn  
 65 70 75 80  
 tgg cct gca aga aaa cat gtt gtg gct tgt gct tcg gtt tgt cgg tct 288  
 Trp Pro Ala Arg Lys His Val Val Ala Cys Ala Ser Val Cys Arg Ser  
 85 90 95

tgg aga gct atg tgc caa gag att gtt ttg ggg cct gaa atc tgt ggg	336
Trp Arg Ala Met Cys Gln Glu Ile Val Leu Gly Pro Glu Ile Cys Gly	
100 105 110	
aaa ctc act ttc cct gtt tcc ctc aaa cag cca ggg cct cgt gat gca	384
Lys Leu Thr Phe Pro Val Ser Leu Lys Gln Pro Gly Pro Arg Asp Ala	
115 120 125	
atg att cag tgt ttc atc aaa agg gat aaa tca aag cta aca ttt cac	432
Met Ile Gln Cys Phe Ile Lys Arg Asp Lys Ser Lys Leu Thr Phe His	
130 135 140	
ctt ttt ctt tgt tta agt ccc gct cta tta gtg gag aat ggg aaa ttt	480
Leu Phe Leu Cys Leu Ser Pro Ala Leu Leu Val Glu Asn Gly Lys Phe	
145 150 155 160	
ctt ctt tca gct aaa aga act cgt aga act act cga acc gag tac att	528
Leu Leu Ser Ala Lys Arg Thr Arg Arg Thr Arg Thr Glu Tyr Ile	
165 170 175	
atc tcc atg gat gct gat aac atc tca aga tcc agc aac tct tac ctc	576
Ile Ser Met Asp Ala Asp Asn Ile Ser Arg Ser Ser Asn Ser Tyr Leu	
180 185 190	
gga aag ctc aga tca aac ttc ctt ggg aca aag ttc ttg gtg tac gac	624
Gly Lys Leu Arg Ser Asn Phe Leu Gly Thr Lys Phe Leu Val Tyr Asp	
195 200 205	
acg caa cca cca cca aac aca tct tcg agc gca ctt atc act gat cga	672
Thr Gln Pro Pro Pro Asn Thr Ser Ser Ser Ala Leu Ile Thr Asp Arg	
210 215 220	
aca agc cga agc agg ttt cac tcc aga cga gtt tct cct aaa gta cca	720
Thr Ser Arg Ser Arg Phe His Ser Arg Arg Val Ser Pro Lys Val Pro	
225 230 235 240	
tcc gga agc tac aac att gct caa atc acc tat gag ctc aac gtg ttg	768
Ser Gly Ser Tyr Asn Ile Ala Gln Ile Thr Tyr Glu Leu Asn Val Leu	
245 250 255	
ggc aca cgc ggg cca cga cga atg cac tgc atc atg aac tcc atc cca	816
Gly Thr Arg Gly Pro Arg Arg Met His Cys Ile Met Asn Ser Ile Pro	
260 265 270	
att tca tcg ctc gaa cca ggc ggt tca gtc cct aac caa ccc gag aaa	864
Ile Ser Ser Leu Glu Pro Gly Gly Ser Val Pro Asn Gln Pro Glu Lys	
275 280 285	
ctc gtc cct gca cca tac tct ctc gac gac tca ttc cgc agt aac atc	912
Leu Val Pro Ala Pro Tyr Ser Leu Asp Asp Ser Phe Arg Ser Asn Ile	
290 295 300	
tcc ttc tcc aaa tca tca ttt gac cac cgc tcc ctc gat ttc agc agt	960
Ser Phe Ser Lys Ser Ser Phe Asp His Arg Ser Leu Asp Phe Ser Ser	
305 310 315 320	

tct aga ttc tcc gaa atg gga ata tcc tgc gac gac aac gaa gaa gaa 1008  
 Ser Arg Phe Ser Glu Met Gly Ile Ser Cys Asp Asp Asn Glu Glu Glu  
                   325                                  330                                  335

gcg agt ttc aga ccg ttg att cta aag aac aag cag cca agg tgg cac 1056  
 Ala Ser Phe Arg Pro Leu Ile Leu Lys Asn Lys Gln Pro Arg Trp His  
                   340                                  345                                  350

gag cag ttg caa tgc tgg tgt ttg aat ttc cgc gga cgt gtg aca gtt 1104  
 Glu Gln Leu Gln Cys Trp Cys Leu Asn Phe Arg Gly Arg Val Thr Val  
                   355                                  360                                  365

gca tcg gtt aag aat ttc cag ctt gta gca gca aga cag ccg cag cct 1152  
 Ala Ser Val Lys Asn Phe Gln Leu Val Ala Ala Arg Gln Pro Gln Pro  
                   370                                  375                                  380

caa ggg aca ggt gca gca gca gca cca aca agt gca cct gct cac cct 1200  
 Gln Gly Thr Gly Ala Ala Ala Ala Pro Thr Ser Ala Pro Ala His Pro  
                   385                                  390                                  395                                  400

gag caa gac aag gtg att ctc cag ttt ggt aaa gta ggg aaa gat atg 1248  
 Glu Gln Asp Lys Val Ile Leu Gln Phe Gly Lys Val Gly Lys Asp Met  
                   405                                  410                                  415

ttc aca atg gac tat agg tat cca tta tcg gcg ttt cag gcg ttt gcg 1296  
 Phe Thr Met Asp Tyr Arg Tyr Pro Leu Ser Ala Phe Gln Ala Phe Ala  
                   420                                  425                                  430

ata tgc tta agc agc ttt gac acc aag ctt gct tgt gaa tag 1338  
 Ile Cys Leu Ser Ser Phe Asp Thr Lys Leu Ala Cys Glu \*  
                   435                                  440                                  445

&lt;210&gt; 130

&lt;211&gt; 445

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 130

Met Ser Phe Arg Gly Ile Val Gln Asp Leu Arg Asp Gly Phe Gly Ser  
 1                  5                  10                  15  
 Leu Ser Arg Arg Ser Phe Asp Phe Arg Leu Ser Ser Leu His Lys Gly  
                   20                  25                  30  
 Lys Ala Gln Gly Ser Ser Phe Arg Glu Tyr Ser Ser Ser Arg Asp Leu  
                   35                  40                  45  
 Leu Ser Pro Val Ile Val Gln Thr Ser Arg Trp Ala Asn Leu Pro Pro  
                   50                  55                  60  
 Glu Leu Leu Phe Asp Val Ile Lys Arg Leu Glu Glu Ser Glu Ser Asn  
 65                  70                  75                  80  
 Trp Pro Ala Arg Lys His Val Val Ala Cys Ala Ser Val Cys Arg Ser  
                   85                  90                  95  
 Trp Arg Ala Met Cys Gln Glu Ile Val Leu Gly Pro Glu Ile Cys Gly  
                   100                  105                  110  
 Lys Leu Thr Phe Pro Val Ser Leu Lys Gln Pro Gly Pro Arg Asp Ala  
                   115                  120                  125  
 Met Ile Gln Cys Phe Ile Lys Arg Asp Lys Ser Lys Leu Thr Phe His

```

      130              135              140
Leu Phe Leu Cys Leu Ser Pro Ala Leu Leu Val Glu Asn Gly Lys Phe
145              150              155              160
Leu Leu Ser Ala Lys Arg Thr Arg Arg Thr Thr Arg Thr Glu Tyr Ile
      165              170              175
Ile Ser Met Asp Ala Asp Asn Ile Ser Arg Ser Ser Asn Ser Tyr Leu
      180              185              190
Gly Lys Leu Arg Ser Asn Phe Leu Gly Thr Lys Phe Leu Val Tyr Asp
      195              200              205
Thr Gln Pro Pro Pro Asn Thr Ser Ser Ser Ala Leu Ile Thr Asp Arg
      210              215              220
Thr Ser Arg Ser Arg Phe His Ser Arg Arg Val Ser Pro Lys Val Pro
225              230              235              240
Ser Gly Ser Tyr Asn Ile Ala Gln Ile Thr Tyr Glu Leu Asn Val Leu
      245              250              255
Gly Thr Arg Gly Pro Arg Arg Met His Cys Ile Met Asn Ser Ile Pro
      260              265              270
Ile Ser Ser Leu Glu Pro Gly Gly Ser Val Pro Asn Gln Pro Glu Lys
      275              280              285
Leu Val Pro Ala Pro Tyr Ser Leu Asp Asp Ser Phe Arg Ser Asn Ile
      290              295              300
Ser Phe Ser Lys Ser Ser Phe Asp His Arg Ser Leu Asp Phe Ser Ser
305              310              315              320
Ser Arg Phe Ser Glu Met Gly Ile Ser Cys Asp Asp Asn Glu Glu Glu
      325              330              335
Ala Ser Phe Arg Pro Leu Ile Leu Lys Asn Lys Gln Pro Arg Trp His
      340              345              350
Glu Gln Leu Gln Cys Trp Cys Leu Asn Phe Arg Gly Arg Val Thr Val
      355              360              365
Ala Ser Val Lys Asn Phe Gln Leu Val Ala Ala Arg Gln Pro Gln Pro
      370              375              380
Gln Gly Thr Gly Ala Ala Ala Ala Pro Thr Ser Ala Pro Ala His Pro
385              390              395              400
Glu Gln Asp Lys Val Ile Leu Gln Phe Gly Lys Val Gly Lys Asp Met
      405              410              415
Phe Thr Met Asp Tyr Arg Tyr Pro Leu Ser Ala Phe Gln Ala Phe Ala
      420              425              430
Ile Cys Leu Ser Ser Phe Asp Thr Lys Leu Ala Cys Glu
      435              440              445

```

<210> 131  
 <211> 1755  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)...(1755)

```

<400> 131
atg gct tac atg tgc act gat agt ggc aat cta atg gct att gct caa   48
Met Ala Tyr Met Cys Thr Asp Ser Gly Asn Leu Met Ala Ile Ala Gln
  1              5              10              15

caa gtc atc aaa cag aag cag caa caa gaa caa caa cag cag caa cat   96
Gln Val Ile Lys Gln Lys Gln Gln Gln Glu Gln Gln Gln Gln His

```

20	25	30	
cat caa gac cat cag att ttt ggt att aat cct ttg tct ctt aac cca	144		
His Gln Asp His Gln Ile Phe Gly Ile Asn Pro Leu Ser Leu Asn Pro			
35 40 45			
tgg ccc aat act tcc ctc ggg ttt ggg ctt tca ggt tcg gct ttt ccc	192		
Trp Pro Asn Thr Ser Leu Gly Phe Gly Leu Ser Gly Ser Ala Phe Pro			
50 55 60			
gac ccg ttt caa gtt acc ggc ggc gga gat tcc aac gat cct ggc ttt	240		
Asp Pro Phe Gln Val Thr Gly Gly Gly Asp Ser Asn Asp Pro Gly Phe			
65 70 75 80			
cct ttt cct aac tta gac cac cac cac gcc aca acc acc ggc ggt ggg	288		
Pro Phe Pro Asn Leu Asp His His His Ala Thr Thr Thr Gly Gly Gly			
85 90 95			
ttc agg tta tct gat ttc ggc ggt gga acc ggc ggc ggc gag ttt gag	336		
Phe Arg Leu Ser Asp Phe Gly Gly Gly Thr Gly Gly Gly Glu Phe Glu			
100 105 110			
tcc gac gag tgg atg gag act ctt atc agc ggt gga gac tcc gtt gca	384		
Ser Asp Glu Trp Met Glu Thr Leu Ile Ser Gly Gly Asp Ser Val Ala			
115 120 125			
gac ggt cct gat tgt gac acc tgg cat gat aat ccc gat tac gta atc	432		
Asp Gly Pro Asp Cys Asp Thr Trp His Asp Asn Pro Asp Tyr Val Ile			
130 135 140			
tac ggt cct gat cca ttc gat act tac ccg agt cga ctc agt gtc caa	480		
Tyr Gly Pro Asp Pro Phe Asp Thr Tyr Pro Ser Arg Leu Ser Val Gln			
145 150 155 160			
ccg tca gat cta aac cga gtc att gac acg tcg agt ccg ctt cct ccg	528		
Pro Ser Asp Leu Asn Arg Val Ile Asp Thr Ser Ser Pro Leu Pro Pro			
165 170 175			
ccg acc ttg tgg cct cct tct tcg cca tta tcg att cct ccg ctt act	576		
Pro Thr Leu Trp Pro Pro Ser Ser Pro Leu Ser Ile Pro Pro Leu Thr			
180 185 190			
cat gag tca cca acc aaa gaa gat cca gag act aac gac tcc gaa gac	624		
His Glu Ser Pro Thr Lys Glu Asp Pro Glu Thr Asn Asp Ser Glu Asp			
195 200 205			
gat gac ttc gac cta gaa cca cct ctc ctc aaa gct ata tac gac tgt	672		
Asp Asp Phe Asp Leu Glu Pro Pro Leu Leu Lys Ala Ile Tyr Asp Cys			
210 215 220			
gca cgg atc tca gac tct gac cct aac gaa gct tcc aag acg ctt ctt	720		
Ala Arg Ile Ser Asp Ser Asp Pro Asn Glu Ala Ser Lys Thr Leu Leu			
225 230 235 240			
cag atc cga gaa tct gta tcg gag cta ggt gat ccg acg gag cga gtt	768		
Gln Ile Arg Glu Ser Val Ser Glu Leu Gly Asp Pro Thr Glu Arg Val			
245 250 255			

gca ttt tac ttc acg gaa gct ctc tcc aac aga ctg tct cct aat tcg	816
Ala Phe Tyr Phe Thr Glu Ala Leu Ser Asn Arg Leu Ser Pro Asn Ser	
260 265 270	
ccg gcg acg tcg tct tct tct tca tct acg gag gat tta atc tta tct	864
Pro Ala Thr Ser Ser Ser Ser Ser Ser Thr Glu Asp Leu Ile Leu Ser	
275 280 285	
tat aaa acc cta aac gac gct tgt cct tac tcc aaa ttc gca cat ttg	912
Tyr Lys Thr Leu Asn Asp Ala Cys Pro Tyr Ser Lys Phe Ala His Leu	
290 295 300	
acg gcg aat caa gcg att cta gaa gcg acg gag aag tcg aac aag att	960
Thr Ala Asn Gln Ala Ile Leu Glu Ala Thr Glu Lys Ser Asn Lys Ile	
305 310 315 320	
cac atc gtc gat ttt gga atc gtt caa ggt ata caa tgg cct gct ctt	1008
His Ile Val Asp Phe Gly Ile Val Gln Gly Ile Gln Trp Pro Ala Leu	
325 330 335	
ctt caa gct cta gct act cgt act tct ggt aaa ccc act caa atc cgg	1056
Leu Gln Ala Leu Ala Thr Arg Thr Ser Gly Lys Pro Thr Gln Ile Arg	
340 345 350	
gtc tcg ggt ata ccc gct cca tct ctc ggt gaa tct ccg gaa ccg tcg	1104
Val Ser Gly Ile Pro Ala Pro Ser Leu Gly Glu Ser Pro Glu Pro Ser	
355 360 365	
tta atc gcc acc gga aac cgc ctc cgt gat ttc gcc aag gtt ctg gat	1152
Leu Ile Ala Thr Gly Asn Arg Leu Arg Asp Phe Ala Lys Val Leu Asp	
370 375 380	
ctg aat ttc gat ttc atc cca att ctc act ccc ata cat tta ctt aac	1200
Leu Asn Phe Asp Phe Ile Pro Ile Leu Thr Pro Ile His Leu Leu Asn	
385 390 395 400	
ggg tca agt ttc cgg gtc gac ccg gat gaa gta ctg gcc gtg aat ttc	1248
Gly Ser Ser Phe Arg Val Asp Pro Asp Glu Val Leu Ala Val Asn Phe	
405 410 415	
atg ctc cag ctc tac aaa tta ctc gac gag acg ccg acg ata gtt gac	1296
Met Leu Gln Leu Tyr Lys Leu Leu Asp Glu Thr Pro Thr Ile Val Asp	
420 425 430	
acc gca cta cgg ctc gcc aaa tcg ttg aac ccg agg gtc gtc act ctc	1344
Thr Ala Leu Arg Leu Ala Lys Ser Leu Asn Pro Arg Val Val Thr Leu	
435 440 445	
gga gaa tac gaa gtg agc tta aac cgg gtc ggt ttc gct aac cgg gta	1392
Gly Glu Tyr Glu Val Ser Leu Asn Arg Val Gly Phe Ala Asn Arg Val	
450 455 460	
aag aac gcg ctt caa ttc tat tcc gcg gtt ttc gaa tcc ctt gaa ccg	1440
Lys Asn Ala Leu Gln Phe Tyr Ser Ala Val Phe Glu Ser Leu Glu Pro	
465 470 475 480	

aac ttg ggg cgt gat tcg gag gag aga gtg aga gtt gag cga gag ttg 1488  
 Asn Leu Gly Arg Asp Ser Glu Glu Arg Val Arg Val Glu Arg Glu Leu  
 485 490 495  
 ttc ggc cgg aga atc tcg ggt ttg att gga ccg gag aaa acc gga att 1536  
 Phe Gly Arg Arg Ile Ser Gly Leu Ile Gly Pro Glu Lys Thr Gly Ile  
 500 505 510  
 cat aga gaa aga atg gaa gag aaa gag caa tgg cgg gta tta atg gag 1584  
 His Arg Glu Arg Met Glu Glu Lys Glu Gln Trp Arg Val Leu Met Glu  
 515 520 525  
 aat gcc ggt ttt gaa tcg gtt aag ctg agt aat tac gca gtg agc caa 1632  
 Asn Ala Gly Phe Glu Ser Val Lys Leu Ser Asn Tyr Ala Val Ser Gln  
 530 535 540  
 gcg aag att cta ttg tgg aat tac aat tac agc aat ttg tat tca att 1680  
 Ala Lys Ile Leu Leu Trp Asn Tyr Asn Tyr Ser Asn Leu Tyr Ser Ile  
 545 550 555 560  
 gtt gaa tct aag cct ggc ttc atc tct ttg gcc tgg aac gat tta cct 1728  
 Val Glu Ser Lys Pro Gly Phe Ile Ser Leu Ala Trp Asn Asp Leu Pro  
 565 570 575  
 ctc ctc act ctt tct tcc tgg cga taa 1755  
 Leu Leu Thr Leu Ser Ser Trp Arg \*  
 580

&lt;210&gt; 132

&lt;211&gt; 584

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (320)...(336)

&lt;223&gt; Conserved domain

&lt;400&gt; 132

Met Ala Tyr Met Cys Thr Asp Ser Gly Asn Leu Met Ala Ile Ala Gln  
 1 5 10 15  
 Gln Val Ile Lys Gln Lys Gln Gln Gln Glu Gln Gln Gln Gln His  
 20 25 30  
 His Gln Asp His Gln Ile Phe Gly Ile Asn Pro Leu Ser Leu Asn Pro  
 35 40 45  
 Trp Pro Asn Thr Ser Leu Gly Phe Gly Leu Ser Gly Ser Ala Phe Pro  
 50 55 60  
 Asp Pro Phe Gln Val Thr Gly Gly Gly Asp Ser Asn Asp Pro Gly Phe  
 65 70 75 80  
 Pro Phe Pro Asn Leu Asp His His His Ala Thr Thr Thr Gly Gly Gly  
 85 90 95  
 Phe Arg Leu Ser Asp Phe Gly Gly Gly Thr Gly Gly Gly Glu Phe Glu  
 100 105 110  
 Ser Asp Glu Trp Met Glu Thr Leu Ile Ser Gly Gly Asp Ser Val Ala  
 115 120 125  
 Asp Gly Pro Asp Cys Asp Thr Trp His Asp Asn Pro Asp Tyr Val Ile



130				135				140							
Tyr	Gly	Pro	Asp	Pro	Phe	Asp	Thr	Tyr	Pro	Ser	Arg	Leu	Ser	Val	Gln
145					150				155						160
Pro	Ser	Asp	Leu	Asn	Arg	Val	Ile	Asp	Thr	Ser	Ser	Pro	Leu	Pro	Pro
				165				170						175	
Pro	Thr	Leu	Trp	Pro	Pro	Ser	Ser	Pro	Leu	Ser	Ile	Pro	Pro	Leu	Thr
				180				185					190		
His	Glu	Ser	Pro	Thr	Lys	Glu	Asp	Pro	Glu	Thr	Asn	Asp	Ser	Glu	Asp
				195			200					205			
Asp	Asp	Phe	Asp	Leu	Glu	Pro	Pro	Leu	Leu	Lys	Ala	Ile	Tyr	Asp	Cys
				210		215					220				
Ala	Arg	Ile	Ser	Asp	Ser	Asp	Pro	Asn	Glu	Ala	Ser	Lys	Thr	Leu	Leu
225					230				235						240
Gln	Ile	Arg	Glu	Ser	Val	Ser	Glu	Leu	Gly	Asp	Pro	Thr	Glu	Arg	Val
				245				250						255	
Ala	Phe	Tyr	Phe	Thr	Glu	Ala	Leu	Ser	Asn	Arg	Leu	Ser	Pro	Asn	Ser
				260				265					270		
Pro	Ala	Thr	Ser	Ser	Ser	Ser	Ser	Ser	Thr	Glu	Asp	Leu	Ile	Leu	Ser
				275			280					285			
Tyr	Lys	Thr	Leu	Asn	Asp	Ala	Cys	Pro	Tyr	Ser	Lys	Phe	Ala	His	Leu
				290		295					300				
Thr	Ala	Asn	Gln	Ala	Ile	Leu	Glu	Ala	Thr	Glu	Lys	Ser	Asn	Lys	Ile
305					310					315					320
His	Ile	Val	Asp	Phe	Gly	Ile	Val	Gln	Gly	Ile	Gln	Trp	Pro	Ala	Leu
				325				330						335	
Leu	Gln	Ala	Leu	Ala	Thr	Arg	Thr	Ser	Gly	Lys	Pro	Thr	Gln	Ile	Arg
				340				345					350		
Val	Ser	Gly	Ile	Pro	Ala	Pro	Ser	Leu	Gly	Glu	Ser	Pro	Glu	Pro	Ser
				355			360					365			
Leu	Ile	Ala	Thr	Gly	Asn	Arg	Leu	Arg	Asp	Phe	Ala	Lys	Val	Leu	Asp
				370		375					380				
Leu	Asn	Phe	Asp	Phe	Ile	Pro	Ile	Leu	Thr	Pro	Ile	His	Leu	Leu	Asn
385					390					395					400
Gly	Ser	Ser	Phe	Arg	Val	Asp	Pro	Asp	Glu	Val	Leu	Ala	Val	Asn	Phe
				405				410						415	
Met	Leu	Gln	Leu	Tyr	Lys	Leu	Leu	Asp	Glu	Thr	Pro	Thr	Ile	Val	Asp
				420				425					430		
Thr	Ala	Leu	Arg	Leu	Ala	Lys	Ser	Leu	Asn	Pro	Arg	Val	Val	Thr	Leu
				435				440				445			
Gly	Glu	Tyr	Glu	Val	Ser	Leu	Asn	Arg	Val	Gly	Phe	Ala	Asn	Arg	Val
				450		455				460					
Lys	Asn	Ala	Leu	Gln	Phe	Tyr	Ser	Al							

<210> 133  
 <211> 1065  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (188)...(760)

<400> 133  
 tcgacccacg cgtccggaca cttacaatt cacaccttct ctttttactc ttcctaaaac 60  
 cctaaatttc ctgccttcag tcttcccact caagtcaacc accaattgaa ttcgatttcg 120  
 aatcattgat ggaaatgatt tgaaaaaaga gtaaagttaa tttttttatt ccttgtaatt 180  
 ttcagaa atg ggg gat tcc gac agg gat tcc ggt gga ggg caa aac ggg 229  
 Met Gly Asp Ser Asp Arg Asp Ser Gly Gly Gly Gln Asn Gly  
 1 5 10

aac aac cag aac gga cag tcc tcc ttg tct cca aga gag caa gac agg 277  
 Asn Asn Gln Asn Gly Gln Ser Ser Leu Ser Pro Arg Glu Gln Asp Arg  
 15 20 25 30

ttc ttg ccg atc gct aac gtc agc cgg atc atg aag aag gcc ttg ccc 325  
 Phe Leu Pro Ile Ala Asn Val Ser Arg Ile Met Lys Lys Ala Leu Pro  
 35 40 45

gcc aac gcc aag atc tct aaa gat gcc aaa gag acg atg cag gag tgt 373  
 Ala Asn Ala Lys Ile Ser Lys Asp Ala Lys Glu Thr Met Gln Glu Cys  
 50 55 60

gtc tcc gag ttc atc agc ttc gtc acc gga gaa gca tct gat aag tgt 421  
 Val Ser Glu Phe Ile Ser Phe Val Thr Gly Glu Ala Ser Asp Lys Cys  
 65 70 75

cag aag gag aag agg aag acg atc aac gga gac gat ttg ctc tgg gct 469  
 Gln Lys Glu Lys Arg Lys Thr Ile Asn Gly Asp Asp Leu Leu Trp Ala  
 80 85 90

atg act act cta ggt ttt gag gat tat gtt gag cca ttg aaa gtt tac 517  
 Met Thr Thr Leu Gly Phe Glu Asp Tyr Val Glu Pro Leu Lys Val Tyr  
 95 100 105 110

ttg cag agg ttt agg gag atc gaa ggg gag agg act gga cta ggg agg 565  
 Leu Gln Arg Phe Arg Glu Ile Glu Gly Glu Arg Thr Gly Leu Gly Arg  
 115 120 125

cca cag act ggt ggt gag gtc gga gag cat cag aga gat gct gtc gga 613  
 Pro Gln Thr Gly Gly Glu Val Gly Glu His Gln Arg Asp Ala Val Gly  
 130 135 140

gat ggc ggt ggg ttc tac ggt ggt ggt ggt ggg atg cag tat cac caa 661  
 Asp Gly Gly Gly Phe Tyr Gly Gly Gly Gly Gly Met Gln Tyr His Gln  
 145 150 155

cat cat cag ttt ctt cac cag cag aac cat atg tat gga gcc aca ggt 709  
 His His Gln Phe Leu His Gln Gln Asn His Met Tyr Gly Ala Thr Gly

160 165 170  
 ggc ggt agc gac agt gga ggt gga gct gcc tcc ggt agg aca agg act 757  
 Gly Gly Ser Asp Ser Gly Gly Gly Ala Ala Ser Gly Arg Thr Arg Thr  
 175 180 185 190  
 taa caaagattgg tgaagtggat ctctctctgt atatagatac ataaatacat 810  
 \*

gtatacacat gcctatTTTT acgacccata taaggatatct atcatgtgat agaacgaaca 870  
 ttggtgttgg tgatgtaaaa tcagatgtgc attaagggtt tagattttga ggctgtgtaa 930  
 aagaagatca agtgtgcttt gttggacaat aggattcact aacgaatctg cttcattgga 990  
 tcttgtatgt aactaaagcc attgtattga atgcaaagt tttcatttgg gatgctttaa 1050  
 aaaaaaaaaa aaaaa 1065

<210> 134  
 <211> 190  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (25)...(116)  
 <223> Conserved domain

<400> 134  
 Met Gly Asp Ser Asp Arg Asp Ser Gly Gly Gly Gln Asn Gly Asn Asn  
 1 5 10 15  
 Gln Asn Gly Gln Ser Ser Leu Ser Pro Arg Glu Gln Asp Arg Phe Leu  
 20 25 30  
 Pro Ile Ala Asn Val Ser Arg Ile Met Lys Lys Ala Leu Pro Ala Asn  
 35 40 45  
 Ala Lys Ile Ser Lys Asp Ala Lys Glu Thr Met Gln Glu Cys Val Ser  
 50 55 60  
 Glu Phe Ile Ser Phe Val Thr Gly Glu Ala Ser Asp Lys Cys Gln Lys  
 65 70 75 80  
 Glu Lys Arg Lys Thr Ile Asn Gly Asp Asp Leu Leu Trp Ala Met Thr  
 85 90 95  
 Thr Leu Gly Phe Glu Asp Tyr Val Glu Pro Leu Lys Val Tyr Leu Gln  
 100 105 110  
 Arg Phe Arg Glu Ile Glu Gly Glu Arg Thr Gly Leu Gly Arg Pro Gln  
 115 120 125  
 Thr Gly Gly Glu Val Gly Glu His Gln Arg Asp Ala Val Gly Asp Gly  
 130 135 140  
 Gly Gly Phe Tyr Gly Gly Gly Gly Met Gln Tyr His Gln His His  
 145 150 155 160  
 Gln Phe Leu His Gln Asn His Met Tyr Gly Ala Thr Gly Gly Gly  
 165 170 175  
 Ser Asp Ser Gly Gly Gly Ala Ala Ser Gly Arg Thr Arg Thr  
 180 185 190

<210> 135  
 <211> 1009  
 <212> DNA  
 <213> Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (27)...(746)

&lt;400&gt; 135

```

gatagtgata acgaaatcct aattccc atg gcc gac aac gac gga gca gtg agt 53
                      Met Ala Asp Asn Asp Gly Ala Val Ser
                        1           5

aac ggc atc ata gtc gag cag acg tca aac aaa gga cct ctt aac gcc 101
Asn Gly Ile Ile Val Glu Gln Thr Ser Asn Lys Gly Pro Leu Asn Ala
 10           15           20           25

gtt aag aaa cca ccg tct aaa gat cga cac agc aaa gtt gac gga aga 149
Val Lys Lys Pro Pro Ser Lys Asp Arg His Ser Lys Val Asp Gly Arg
           30           35           40

gga aga agg att cgt atg cca atc att tgc gca gct cga gtt ttt caa 197
Gly Arg Arg Ile Arg Met Pro Ile Ile Cys Ala Ala Arg Val Phe Gln
           45           50           55

ttg acc aga gag tta ggt cac aag tcc gat ggt caa acc ata gag tgg 245
Leu Thr Arg Glu Leu Gly His Lys Ser Asp Gly Gln Thr Ile Glu Trp
           60           65           70

ctt ctc cgt caa gct gag cct tct atc ata gcc gcc act gga act ggc 293
Leu Leu Arg Gln Ala Glu Pro Ser Ile Ile Ala Ala Thr Gly Thr Gly
           75           80           85

act act ccg gcg agt ttc tcc act gct tct ctc tcc act tct tct ccg 341
Thr Thr Pro Ala Ser Phe Ser Thr Ala Ser Leu Ser Thr Ser Ser Pro
           90           95          100          105

ttt act ctc ggg aaa cgt gtc gtc aga gcg gag gaa gga gaa tcc ggc 389
Phe Thr Leu Gly Lys Arg Val Val Arg Ala Glu Glu Gly Glu Ser Gly
           110          115          120

ggc gga gga gga gga ggg tta aca gtg gga cac aca atg ggg act tcg 437
Gly Gly Gly Gly Gly Gly Leu Thr Val Gly His Thr Met Gly Thr Ser
           125          130          135

tta atg ggt ggt ggt ggt tct ggt ggg ttt tgg gct gtt ccg gcg agg 485
Leu Met Gly Gly Gly Gly Ser Gly Gly Phe Trp Ala Val Pro Ala Arg
           140          145          150

ccg gat ttc gga caa gtc tgg agc ttt gca acc gga gct cca ccg gaa 533
Pro Asp Phe Gly Gln Val Trp Ser Phe Ala Thr Gly Ala Pro Pro Glu
           155          160          165

atg gtt ttt gcg cag cag cag caa cca gct aca ctc ttc gtc cgc cac 581
Met Val Phe Ala Gln Gln Gln Gln Pro Ala Thr Leu Phe Val Arg His
           170          175          180          185

cag cag caa cag caa gct tcc gcc gcc gca gca gct gca atg ggt gag 629
Gln Gln Gln Gln Gln Ala Ser Ala Ala Ala Ala Met Gly Glu
           190          195          200

```

gct tca gca gct aga gtt ggg aat tat ctt ccg ggt cat cat ctc aat 677  
 Ala Ser Ala Ala Arg Val Gly Asn Tyr Leu Pro Gly His His Leu Asn  
 205 210 215

ttg ctt gct tct ttg tct ggt gga gct aac ggg tcg ggt cgg agg gaa 725  
 Leu Leu Ala Ser Leu Ser Gly Gly Ala Asn Gly Ser Gly Arg Arg Glu  
 220 225 230

gac gac cac gaa cca cgt tga gaaatgggtat tgtctttttg gtaatgtata 776  
 Asp Asp His Glu Pro Arg \*  
 235

gaaaaattcc tatgttttat gtcacgcgaaa gtgttttagaa agtacctcta atttgcgggt 836  
 tcttttgctc cttttttact taatttaagc ttattgcttg ttgattagg gttttagggt 896  
 ttaagaatat ttggtctcgt taatttggtt cggagagtga tagaaagaga gagagattga 956  
 ttgattgttg tacctaaaac gctataaaag ctctgttttt actagcgaaa aaa 1009

<210> 136

<211> 239

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (32)...(93)

<223> Conserved domain

<400> 136

Met Ala Asp Asn Asp Gly Ala Val Ser Asn Gly Ile Ile Val Glu Gln  
 1 5 10 15  
 Thr Ser Asn Lys Gly Pro Leu Asn Ala Val Lys Lys Pro Pro Ser Lys  
 20 25 30  
 Asp Arg His Ser Lys Val Asp Gly Arg Gly Arg Arg Ile Arg Met Pro  
 35 40 45  
 Ile Ile Cys Ala Ala Arg Val Phe Gln Leu Thr Arg Glu Leu Gly His  
 50 55 60  
 Lys Ser Asp Gly Gln Thr Ile Glu Trp Leu Leu Arg Gln Ala Glu Pro  
 65 70 75 80  
 Ser Ile Ile Ala Ala Thr Gly Thr Gly Thr Thr Pro Ala Ser Phe Ser  
 85 90 95  
 Thr Ala Ser Leu Ser Thr Ser Ser Pro Phe Thr Leu Gly Lys Arg Val  
 100 105 110  
 Val Arg Ala Glu Glu Gly Glu Ser Gly Gly Gly Gly Gly Gly Gly Leu  
 115 120 125  
 Thr Val Gly His Thr Met Gly Thr Ser Leu Met Gly Gly Gly Gly Ser  
 130 135 140  
 Gly Gly Phe Trp Ala Val Pro Ala Arg Pro Asp Phe Gly Gln Val Trp  
 145 150 155 160  
 Ser Phe Ala Thr Gly Ala Pro Pro Glu Met Val Phe Ala Gln Gln Gln  
 165 170 175  
 Gln Pro Ala Thr Leu Phe Val Arg His Gln Gln Gln Gln Gln Ala Ser  
 180 185 190  
 Ala Ala Ala Ala Ala Met Gly Glu Ala Ser Ala Ala Arg Val Gly  
 195 200 205  
 Asn Tyr Leu Pro Gly His His Leu Asn Leu Leu Ala Ser Leu Ser Gly  
 210 215 220

Gly Ala Asn Gly Ser Gly Arg Arg Glu Asp Asp His Glu Pro Arg  
 225 230 235

<210> 137

<211> 1281

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (64)...(1098)

<400> 137

```

cacaacacaa acacatttct gttttctcca ttgtttcaaa ccataaaaaa aaacacagat 60
taa atg gaa tcg agt agc gtt gat gag agt act aca agt aca ggt tcc 108
  Met Glu Ser Ser Ser Val Asp Glu Ser Thr Thr Ser Thr Gly Ser
    1          5          10          15

atc tgt gaa acc ccg gcg ata act ccg gcg aaa aag tcg tcg gta ggt 156
Ile Cys Glu Thr Pro Ala Ile Thr Pro Ala Lys Lys Ser Ser Val Gly
          20          25          30

aac tta tac agg atg gga agc gga tca agc gtt gtg tta gat tca gag 204
Asn Leu Tyr Arg Met Gly Ser Gly Ser Ser Val Val Leu Asp Ser Glu
          35          40          45

aac ggc gta gaa gct gaa tct agg aag ctt ccg tcg tca aaa tac aaa 252
Asn Gly Val Glu Ala Glu Ser Arg Lys Leu Pro Ser Ser Lys Tyr Lys
          50          55          60

ggt gtg gtg cca caa cca aac gga aga tgg gga gct cag att tac gag 300
Gly Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu
          65          70          75

aaa cac cag cgc gtg tgg ctc ggg aca ttc aac gaa gaa gac gaa gcc 348
Lys His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Glu Asp Glu Ala
          80          85          90          95

gct cgt gcc tac gac gtc gcg gtt cac agg ttc cgt cgc cgt gac gcc 396
Ala Arg Ala Tyr Asp Val Ala Val His Arg Phe Arg Arg Arg Asp Ala
          100          105          110

gtc aca aat ttc aaa gac gtg aag atg gac gaa gac gag gtc gat ttc 444
Val Thr Asn Phe Lys Asp Val Lys Met Asp Glu Asp Glu Val Asp Phe
          115          120          125

ttg aat tct cat tcg aaa tct gag atc gtt gat atg ttg agg aaa cat 492
Leu Asn Ser His Ser Lys Ser Glu Ile Val Asp Met Leu Arg Lys His
          130          135          140

act tat aac gaa gag tta gag cag agt aaa cgg cgt cgt aat ggt aac 540
Thr Tyr Asn Glu Glu Leu Glu Gln Ser Lys Arg Arg Arg Asn Gly Asn
          145          150          155

gga aac atg act agg acg ttg tta acg tcg ggg ttg agt aat gat ggt 588
Gly Asn Met Thr Arg Thr Leu Leu Thr Ser Gly Leu Ser Asn Asp Gly

```

```

160          165          170          175
gtt tct acg acg ggg ttt aga tcg gcg gag gca ctg ttt gag aaa gcg 636
Val Ser Thr Thr Gly Phe Arg Ser Ala Glu Ala Leu Phe Glu Lys Ala
          180          185          190

gta acg cca agc gac gtt ggg aag cta aac cgt ttg gtt ata ccg aaa 684
Val Thr Pro Ser Asp Val Gly Lys Leu Asn Arg Leu Val Ile Pro Lys
          195          200          205

cat cac gca gag aaa cat ttt ccg tta ccg tca agt aac gtt tcc gtg 732
His His Ala Glu Lys His Phe Pro Leu Pro Ser Ser Asn Val Ser Val
          210          215          220

aaa gga gtg ttg ttg aac ttt gag gac gtt aac ggg aaa gtg tgg agg 780
Lys Gly Val Leu Leu Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg
          225          230          235

ttc cgt tac tcg tat tgg aac agt agt cag agt tat gtt ttg act aaa 828
Phe Arg Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys
          240          245          250          255

ggt tgg agc agg ttc gtt aag gag aag aat cta cgt gct ggt gac gtg 876
Gly Trp Ser Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val
          260          265          270

gtt agt ttc agt aga tct aac ggt cag gat caa cag ttg tac att ggg 924
Val Ser Phe Ser Arg Ser Asn Gly Gln Asp Gln Gln Leu Tyr Ile Gly
          275          280          285

tgg aag tcg aga tcc ggg tca gat tta gat gcg ggt cgg gtt ttg aga 972
Trp Lys Ser Arg Ser Gly Ser Asp Leu Asp Ala Gly Arg Val Leu Arg
          290          295          300

ttg ttc gga gtt aac att tca ccg gag agt tca aga aac gac gtc gta 1020
Leu Phe Gly Val Asn Ile Ser Pro Glu Ser Ser Arg Asn Asp Val Val
          305          310          315

gga aac aaa aga gtg aac gat act gag atg tta tcg ttg gtg tgt agc 1068
Gly Asn Lys Arg Val Asn Asp Thr Glu Met Leu Ser Leu Val Cys Ser
          320          325          330          335

aag aag caa cgc atc ttt cac gcc tcg taa caactcttct tctttttttt 1118
Lys Lys Gln Arg Ile Phe His Ala Ser *
          340

tcttttggtg ttttaataat ttttaaaaac tccattttcg ttttctttat ttgcacgggt 1178
tctttttctt ttgtttacca aaggttcatg agttgttttt gttgtattga tgaactgtaa 1238
atttttattta taggataaat ttttaaaaaa aaaaaaaaaa aaa 1281

```

&lt;210&gt; 138

&lt;211&gt; 344

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (59)...(124)

&lt;223&gt; Conserved domain

&lt;400&gt; 138

```

Met Glu Ser Ser Ser Val Asp Glu Ser Thr Thr Ser Thr Gly Ser Ile
 1          5          10          15
Cys Glu Thr Pro Ala Ile Thr Pro Ala Lys Lys Ser Ser Val Gly Asn
          20          25          30
Leu Tyr Arg Met Gly Ser Gly Ser Ser Val Val Leu Asp Ser Glu Asn
          35          40          45
Gly Val Glu Ala Glu Ser Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly
          50          55          60
Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys
65          70          75          80
His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Glu Asp Glu Ala Ala
          85          90          95
Arg Ala Tyr Asp Val Ala Val His Arg Phe Arg Arg Arg Asp Ala Val
          100          105          110
Thr Asn Phe Lys Asp Val Lys Met Asp Glu Asp Glu Val Asp Phe Leu
          115          120          125
Asn Ser His Ser Lys Ser Glu Ile Val Asp Met Leu Arg Lys His Thr
130          135          140
Tyr Asn Glu Glu Leu Glu Gln Ser Lys Arg Arg Arg Asn Gly Asn Gly
145          150          155          160
Asn Met Thr Arg Thr Leu Leu Thr Ser Gly Leu Ser Asn Asp Gly Val
          165          170          175
Ser Thr Thr Gly Phe Arg Ser Ala Glu Ala Leu Phe Glu Lys Ala Val
          180          185          190
Thr Pro Ser Asp Val Gly Lys Leu Asn Arg Leu Val Ile Pro Lys His
          195          200          205
His Ala Glu Lys His Phe Pro Leu Pro Ser Ser Asn Val Ser Val Lys
210          215          220
Gly Val Leu Leu Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg Phe
225          230          235          240
Arg Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly
          245          250          255
Trp Ser Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val
          260          265          270
Ser Phe Ser Arg Ser Asn Gly Gln Asp Gln Gln Leu Tyr Ile Gly Trp
          275          280          285
Lys Ser Arg Ser Gly Ser Asp Leu Asp Ala Gly Arg Val Leu Arg Leu
290          295          300
Phe Gly Val Asn Ile Ser Pro Glu Ser Ser Arg Asn Asp Val Val Gly
305          310          315          320
Asn Lys Arg Val Asn Asp Thr Glu Met Leu Ser Leu Val Cys Ser Lys
          325          330          335
Lys Gln Arg Ile Phe His Ala Ser
          340

```

&lt;210&gt; 139

&lt;211&gt; 1879

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS



&lt;222&gt; (31)...(1575)

&lt;400&gt; 139

```

tttttttttg tttgttaatt ttggggatcg atg tcg gaa aag gaa gaa gct ccg 54
                        Met Ser Glu Lys Glu Glu Ala Pro
                          1                      5

tcg aca tcg aag tcc acc gga gct ccg tcg cgt ccg act tta tct ctt 102
Ser Thr Ser Lys Ser Thr Gly Ala Pro Ser Arg Pro Thr Leu Ser Leu
   10                      15                      20

cct cca cgg ccg ttt agt gag atg ttc ttt aac ggt ggc gtt gga ttc 150
Pro Pro Arg Pro Phe Ser Glu Met Phe Phe Asn Gly Gly Val Gly Phe
   25                      30                      35                      40

agt cct ggt ccg atg act ctg gtc tct aat atg ttc cct gat tcc gat 198
Ser Pro Gly Pro Met Thr Leu Val Ser Asn Met Phe Pro Asp Ser Asp
                      45                      50                      55

gag ttt agg tct ttc tct cag ctt ctg gct gga gcc atg tct tct cca 246
Glu Phe Arg Ser Phe Ser Gln Leu Leu Ala Gly Ala Met Ser Ser Pro
                      60                      65                      70

gcg act gca gct gct gct gct gct gct gcg acg gct agt gat tac cag 294
Ala Thr Ala Ala Ala Ala Ala Ala Ala Thr Ala Ser Asp Tyr Gln
   75                      80                      85

aga ctt ggt gaa ggg act aat agc tct agt ggt gat gtt gac ccg aga 342
Arg Leu Gly Glu Gly Thr Asn Ser Ser Ser Gly Asp Val Asp Pro Arg
   90                      95                      100

ttc aag caa aac aga cca acc ggt ttg atg att tct caa tct caa tcg 390
Phe Lys Gln Asn Arg Pro Thr Gly Leu Met Ile Ser Gln Ser Gln Ser
  105                      110                      115                      120

ccg tcg atg ttc acc gta ccg cct ggt tta agt cca gct atg ttg ctg 438
Pro Ser Met Phe Thr Val Pro Pro Gly Leu Ser Pro Ala Met Leu Leu
                      125                      130                      135

gat tca cca agc ttt ttg ggt ctt ttc tct ccc gtt cag gga tca tat 486
Asp Ser Pro Ser Phe Leu Gly Leu Phe Ser Pro Val Gln Gly Ser Tyr
                      140                      145                      150

gga atg aca cat cag caa gct cta gct caa gtc act gct caa gca gtt 534
Gly Met Thr His Gln Gln Ala Leu Ala Gln Val Thr Ala Gln Ala Val
                      155                      160                      165

caa gcc aat gcc aat atg caa cca caa aca gag tac cct cct ccc tct 582
Gln Ala Asn Ala Asn Met Gln Pro Gln Thr Glu Tyr Pro Pro Pro Ser
                      170                      175                      180

caa gtt caa tca ttt tca tcg ggt caa gcg cag atc ccg acc tcg gct 630
Gln Val Gln Ser Phe Ser Ser Gly Gln Ala Gln Ile Pro Thr Ser Ala
  185                      190                      195                      200

cca cta cca gct caa aga gaa acc tca gat gta acc atc ata gag cac 678
Pro Leu Pro Ala Gln Arg Glu Thr Ser Asp Val Thr Ile Ile Glu His

```

205	210	215	
agg tca caa cag cct cta aat gtt gac aaa cca gct gat gat ggc tat 726			
Arg Ser Gln Gln Pro Leu Asn Val Asp Lys Pro Ala Asp Asp Gly Tyr			
220	225	230	
aac tgg cga aaa tat ggg caa aag caa gtt aaa ggt agc gag ttt cca 774			
Asn Trp Arg Lys Tyr Gly Gln Lys Gln Val Lys Gly Ser Glu Phe Pro			
235	240	245	
cga agc tat tac aag tgt act aat cca gga tgt cct gtc aag aag aag 822			
Arg Ser Tyr Tyr Lys Cys Thr Asn Pro Gly Cys Pro Val Lys Lys Lys			
250	255	260	
gtt gag aga tct ctt gat gga caa gta acg gag att atc tac aaa ggt 870			
Val Glu Arg Ser Leu Asp Gly Gln Val Thr Glu Ile Ile Tyr Lys Gly			
265	270	275	280
cag cac aat cat gaa cct cct caa aac act aag cga ggt aac aaa gat 918			
Gln His Asn His Glu Pro Pro Gln Asn Thr Lys Arg Gly Asn Lys Asp			
285	290	295	
aac acc gcg aat ata aat ggg agt tcg ata aat aac aat cgc ggg agt 966			
Asn Thr Ala Asn Ile Asn Gly Ser Ser Ile Asn Asn Asn Arg Gly Ser			
300	305	310	
tct gaa ttg ggg gca tca cag ttt caa act aat agc tcc aac aag act 1014			
Ser Glu Leu Gly Ala Ser Gln Phe Gln Thr Asn Ser Ser Asn Lys Thr			
315	320	325	
aag aga gag caa cat gaa gca gta agt caa gct acg aca aca gag cac 1062			
Lys Arg Glu Gln His Glu Ala Val Ser Gln Ala Thr Thr Thr Glu His			
330	335	340	
ttg tct gag gca agt gac ggt gaa gaa gtt ggt aat gga gaa act gat 1110			
Leu Ser Glu Ala Ser Asp Gly Glu Glu Val Gly Asn Gly Glu Thr Asp			
345	350	355	360
gtg aga gag aaa gat gag aat gag cct gat ccc aag aga aga agt aca 1158			
Val Arg Glu Lys Asp Glu Asn Glu Pro Asp Pro Lys Arg Arg Ser Thr			
365	370	375	
gaa gtt cgg att tca gaa cca gct cct gct gct tca cat aga act gtg 1206			
Glu Val Arg Ile Ser Glu Pro Ala Pro Ala Ala Ser His Arg Thr Val			
380	385	390	
aca gag cct aga att att gtc caa acg acg agt gaa gtt gat ctt cta 1254			
Thr Glu Pro Arg Ile Ile Val Gln Thr Thr Ser Glu Val Asp Leu Leu			
395	400	405	
gat gat gga tat agg tgg cgt aaa tat gga cag aaa gtt gtc aaa ggg 1302			
Asp Asp Gly Tyr Arg Trp Arg Lys Tyr Gly Gln Lys Val Val Lys Gly			
410	415	420	
aat cct tat ccg agg agc tac tac aag tgc aca aca cca gga tgt ggt 1350			
Asn Pro Tyr Pro Arg Ser Tyr Tyr Lys Cys Thr Thr Pro Gly Cys Gly			
425	430	435	440

gtg agg aaa cat gta gag aga gca gca aca gat cca aaa gct gta gta 1398  
 Val Arg Lys His Val Glu Arg Ala Ala Thr Asp Pro Lys Ala Val Val  
                   445                                  450                                  455

aca aca tat gaa gga aaa cat aac cat gac ctt ccc gct gct aaa tca 1446  
 Thr Thr Tyr Glu Gly Lys His Asn His Asp Leu Pro Ala Ala Lys Ser  
                   460                                  465                                  470

agc agc cat gcc gct gca gcg gca cag tta agg cca gat aat cga cct 1494  
 Ser Ser His Ala Ala Ala Ala Ala Gln Leu Arg Pro Asp Asn Arg Pro  
                   475                                  480                                  485

ggc ggt ttg gct aac tta aat caa cag cag cag caa cag ccc gtt gcg 1542  
 Gly Gly Leu Ala Asn Leu Asn Gln Gln Gln Gln Gln Gln Pro Val Ala  
                   490                                  495                                  500

cgg cta agg ctt aaa gaa gag caa aca act tga gagaagaaaa ctcttgaccg 1595  
 Arg Leu Arg Leu Lys Glu Glu Thr Thr \*  
 505                                  510

tttttcatta caaaagcttt caaattccac tcacacactt gtctgaaaaa tctagcagtt 1655  
 tgcaggaaag aaacagcttc aagagggtgt agttcttcta tggtctggtg taaaacttaa 1715  
 aagcttttta gggttttcag atttctgttt actaatactg tatgtgaatt cttttgtaca 1775  
 tgaggaagaa aattacaggg ggatattttg tggtgtatct tttgtgttat tgtttcagta 1835  
 aaagataggt cttacatttt gtgtaaaaaa aaaaaaaaaa aaaa 1879

<210> 140

<211> 514

<212> PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (227)...(285)

<223> Conserved domain

<221> DOMAIN

<222> (407)...(465)

<223> Conserved domain

<400> 140

Met Ser Glu Lys Glu Glu Ala Pro Ser Thr Ser Lys Ser Thr Gly Ala  
 1                                  5                                  10                                  15  
 Pro Ser Arg Pro Thr Leu Ser Leu Pro Pro Arg Pro Phe Ser Glu Met  
                   20                                  25                                  30  
 Phe Phe Asn Gly Gly Val Gly Phe Ser Pro Gly Pro Met Thr Leu Val  
                   35                                  40                                  45  
 Ser Asn Met Phe Pro Asp Ser Asp Glu Phe Arg Ser Phe Ser Gln Leu  
                   50                                  55                                  60  
 Leu Ala Gly Ala Met Ser Ser Pro Ala Thr Ala Ala Ala Ala Ala Ala  
 65                                  70                                  75                                  80  
 Ala Ala Thr Ala Ser Asp Tyr Gln Arg Leu Gly Glu Gly Thr Asn Ser  
                   85                                  90                                  95  
 Ser Ser Gly Asp Val Asp Pro Arg Phe Lys Gln Asn Arg Pro Thr Gly  
                   100                                  105                                  110  
 Leu Met Ile Ser Gln Ser Gln Ser Pro Ser Met Phe Thr Val Pro Pro

115	120	125
Gly Leu Ser Pro Ala Met	Leu Leu Asp Ser Pro Ser	Phe Leu Gly Leu
130	135	140
Phe Ser Pro Val Gln Gly	Ser Tyr Gly Met Thr His	Gln Gln Ala Leu
145	150	155
Ala Gln Val Thr Ala Gln	Ala Val Gln Ala Asn Ala	Asn Met Gln Pro
165	170	175
Gln Thr Glu Tyr Pro Pro	Pro Ser Gln Val Gln Ser	Phe Ser Ser Gly
180	185	190
Gln Ala Gln Ile Pro Thr	Ser Ala Pro Leu Pro Ala	Gln Arg Glu Thr
195	200	205
Ser Asp Val Thr Ile Ile	Glu His Arg Ser Gln Gln	Pro Leu Asn Val
210	215	220
Asp Lys Pro Ala Asp Asp	Gly Tyr Asn Trp Arg Lys	Tyr Gly Gln Lys
225	230	235
Gln Val Lys Gly Ser Glu	Phe Pro Arg Ser Tyr Tyr	Lys Cys Thr Asn
245	250	255
Pro Gly Cys Pro Val Lys	Lys Lys Val Glu Arg Ser	Leu Asp Gly Gln
260	265	270
Val Thr Glu Ile Ile Tyr	Lys Gly Gln His Asn His	Glu Pro Pro Gln
275	280	285
Asn Thr Lys Arg Gly Asn	Lys Asp Asn Thr Ala Asn	Ile Asn Gly Ser
290	295	300
Ser Ile Asn Asn Asn Arg	Gly Ser Ser Glu Leu Gly	Ala Ser Gln Phe
305	310	315
Gln Thr Asn Ser Ser Asn	Lys Thr Lys Arg Glu Gln	His Glu Ala Val
325	330	335
Ser Gln Ala Thr Thr Thr	Glu His Leu Ser Glu Ala	Ser Asp Gly Glu
340	345	350
Glu Val Gly Asn Gly Glu	Thr Asp Val Arg Glu Lys	Asp Glu Asn Glu
355	360	365
Pro Asp Pro Lys Arg Arg	Ser Thr Glu Val Arg Ile	Ser Glu Pro Ala
370	375	380
Pro Ala Ala Ser His Arg	Thr Val Thr Glu Pro Arg	Ile Ile Val Gln
385	390	395
Thr Thr Ser Glu Val Asp	Leu Leu Asp Asp Gly Tyr	Arg Trp Arg Lys
405	410	415
Tyr Gly Gln Lys Val Val	Lys Gly Asn Pro Tyr Pro	Arg Ser Tyr Tyr
420	425	430
Lys Cys Thr Thr Pro Gly	Cys Gly Val Arg Lys His	Val Glu Arg Ala
435	440	445
Ala Thr Asp Pro Lys Ala	Val Val Thr Thr Tyr Glu	Gly Lys His Asn
450	455	460
His Asp Leu Pro Ala Ala	Lys Ser Ser Ser His Ala	Ala Ala Ala Ala
465	470	475
Gln Leu Arg Pro Asp Asn	Arg Pro Gly Gly Leu Ala	Asn Leu Asn Gln
485	490	495
Gln Gln Gln Gln Gln Pro	Val Ala Arg Leu Arg Leu	Lys Glu Glu Gln
500	505	510
Thr Thr		

&lt;210&gt; 141

&lt;211&gt; 485

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (26)...(457)

&lt;400&gt; 141

```

atctcatctt catccacca aaaac atg gat tca aga gac acc gga gaa act   52
                        Met Asp Ser Arg Asp Thr Gly Glu Thr
                        1           5

gac cag agc aag tac aaa ggt atc cgt cgt cgg aaa tgg gga aaa tgg   100
Asp Gln Ser Lys Tyr Lys Gly Ile Arg Arg Arg Lys Trp Gly Lys Trp
10           15           20           25

gta tca gag att cgt gtc ccg gga act cgt caa cgt ctc tgg tta ggc   148
Val Ser Glu Ile Arg Val Pro Gly Thr Arg Gln Arg Leu Trp Leu Gly
           30           35           40

tct ttc tcc acc gca gaa ggc gct gcc gta gcc cac gac gtc gct ttt   196
Ser Phe Ser Thr Ala Glu Gly Ala Val Ala His Asp Val Ala Phe
           45           50           55

tac tgc ttg cac cga cca tct tcc ctc gac gac gaa tct ttt aac ttc   244
Tyr Cys Leu His Arg Pro Ser Ser Leu Asp Asp Glu Ser Phe Asn Phe
           60           65           70

cct cac tta ctt aca acc tcc ctc gcc tcc aat ata tct cct aag tcc   292
Pro His Leu Leu Thr Thr Ser Leu Ala Ser Asn Ile Ser Pro Lys Ser
           75           80           85

atc caa aaa gct gct tcc gac gcc ggc atg gcc gtg gac gcc gga ttc   340
Ile Gln Lys Ala Ala Ser Asp Ala Gly Met Ala Val Asp Ala Gly Phe
           90           95           100           105

cat ggt gct gtg tct ggg agt ggt ggt tgt gaa gag aga tct tcc atg   388
His Gly Ala Val Ser Gly Ser Gly Gly Cys Glu Glu Arg Ser Ser Met
           110           115           120

gcg aat atg gag gag gag gac aaa ctt agt atc tcc gtg tat gat tat   436
Ala Asn Met Glu Glu Glu Asp Lys Leu Ser Ile Ser Val Tyr Asp Tyr
           125           130           135

ctt gaa gac gat ctc gtt tga tctatacgag tacgttttta gcagttaa   485
Leu Glu Asp Asp Leu Val *
           140

```

&lt;210&gt; 142

&lt;211&gt; 143

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (11)...(83)

&lt;223&gt; Conserved domain

&lt;400&gt; 142

```

Met Asp Ser Arg Asp Thr Gly Glu Thr Asp Gln Ser Lys Tyr Lys Gly
 1           5           10           15
Ile Arg Arg Arg Lys Trp Gly Lys Trp Val Ser Glu Ile Arg Val Pro
          20           25           30
Gly Thr Arg Gln Arg Leu Trp Leu Gly Ser Phe Ser Thr Ala Glu Gly
      35           40           45
Ala Ala Val Ala His Asp Val Ala Phe Tyr Cys Leu His Arg Pro Ser
      50           55           60
Ser Leu Asp Asp Glu Ser Phe Asn Phe Pro His Leu Leu Thr Thr Ser
      65           70           75           80
Leu Ala Ser Asn Ile Ser Pro Lys Ser Ile Gln Lys Ala Ala Ser Asp
          85           90           95
Ala Gly Met Ala Val Asp Ala Gly Phe His Gly Ala Val Ser Gly Ser
          100          105          110
Gly Gly Cys Glu Glu Arg Ser Ser Met Ala Asn Met Glu Glu Glu Asp
          115          120          125
Lys Leu Ser Ile Ser Val Tyr Asp Tyr Leu Glu Asp Asp Leu Val
      130           135           140

```

&lt;210&gt; 143

&lt;211&gt; 931

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (66)...(899)

&lt;400&gt; 143

```

tcctcactct ctctcttttt ctctaaccat aaaatctctt tgatctcttt ctctgtgttt 60
tgata atg gaa aat gtt ggt gtt ggg atg ccg ttt tac gat tta ggg caa 110
      Met Glu Asn Val Gly Val Gly Met Pro Phe Tyr Asp Leu Gly Gln
        1           5           10           15

aca agg gtt tac cca ctc ttg tct gat ttc cac gat tta tcg gcg gag 158
Thr Arg Val Tyr Pro Leu Leu Ser Asp Phe His Asp Leu Ser Ala Glu
          20           25           30

agg tat ccg gta ggg ttc atg gat tta ctg ggt gtt cat cgt cat aca 206
Arg Tyr Pro Val Gly Phe Met Asp Leu Leu Gly Val His Arg His Thr
          35           40           45

ccc acc cat acg ccg ttg atg cat ttt ccg acc aca cct aac tcg tcc 254
Pro Thr His Thr Pro Leu Met His Phe Pro Thr Thr Pro Asn Ser Ser
          50           55           60

tcg agc gaa gct gtg aat gga gat gac gaa gaa gaa gaa gat gga gaa 302
Ser Ser Glu Ala Val Asn Gly Asp Asp Glu Glu Glu Glu Asp Gly Glu
          65           70           75

gaa cag cag cat aag aca aag aag cgg ttt aaa ttc act aaa atg agt 350
Glu Gln Gln His Lys Thr Lys Lys Arg Phe Lys Phe Thr Lys Met Ser
          80           85           90           95

aga aag cag acg aag aag aag gtg cca aaa gtg tca ttc atc acg agg 398

```

Arg Lys Gln Thr Lys Lys Lys Val Pro Lys Val Ser Phe Ile Thr Arg  
 100 105 110  
 agt gag gtt ctt cat cta gat gat ggt tat aag tgg aga aaa tac ggt 446  
 Ser Glu Val Leu His Leu Asp Asp Gly Tyr Lys Trp Arg Lys Tyr Gly  
 115 120 125  
 caa aaa cct gtc aaa gac agc cct ttt cca aga aat tat tac cgt tgc 494  
 Gln Lys Pro Val Lys Asp Ser Pro Phe Pro Arg Asn Tyr Tyr Arg Cys  
 130 135 140  
 aca aca act tgg tgt gac gtg aag aag aga gta gag aga tca ttc agt 542  
 Thr Thr Thr Trp Cys Asp Val Lys Lys Arg Val Glu Arg Ser Phe Ser  
 145 150 155  
 gat cca agc agt gta atc acc act tac gaa ggt caa cat act cat cct 590  
 Asp Pro Ser Ser Val Ile Thr Thr Tyr Glu Gly Gln His Thr His Pro  
 160 165 170 175  
 cgt cca cta ctc atc atg ccc aaa gaa ggc agc tct cca tcc aat ggc 638  
 Arg Pro Leu Leu Ile Met Pro Lys Glu Gly Ser Ser Pro Ser Asn Gly  
 180 185 190  
 tca gct tct agg gcc cac att ggc ctc cct aca ctc cct cct cag ctt 686  
 Ser Ala Ser Arg Ala His Ile Gly Leu Pro Thr Leu Pro Pro Gln Leu  
 195 200 205  
 tta gat tac aac aac caa caa caa caa gcg ccg tct tct ttt gga acc 734  
 Leu Asp Tyr Asn Asn Gln Gln Gln Ala Pro Ser Ser Phe Gly Thr  
 210 215 220  
 gag tac att aac agg caa gaa aaa gga att aat cat gat gat gat gac 782  
 Glu Tyr Ile Asn Arg Gln Glu Lys Gly Ile Asn His Asp Asp Asp  
 225 230 235  
 gat cat gtt gtg aag aag agt cga act cgg gat ctg ctg gat gga gct 830  
 Asp His Val Val Lys Lys Ser Arg Thr Arg Asp Leu Leu Asp Gly Ala  
 240 245 250 255  
 ggt tta gtc aaa gat cat ggc ctt ctt cag gat gtt gtt ccc tct cat 878  
 Gly Leu Val Lys Asp His Gly Leu Leu Gln Asp Val Val Pro Ser His  
 260 265 270  
 atc att aag gaa gag tat tag ttaatcgcat aattatgtag ctagctagct 929  
 Ile Ile Lys Glu Glu Tyr \*  
 275  
 ag 931  
 <210> 144  
 <211> 277  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 144  
 Met Glu Asn Val Gly Val Gly Met Pro Phe Tyr Asp Leu Gly Gln Thr  
 1 5 10 15

Arg Val Tyr Pro Leu Leu Ser Asp Phe His Asp Leu Ser Ala Glu Arg  
                   20                  25                  30  
 Tyr Pro Val Gly Phe Met Asp Leu Leu Gly Val His Arg His Thr Pro  
           35                  40                  45  
 Thr His Thr Pro Leu Met His Phe Pro Thr Thr Pro Asn Ser Ser Ser  
       50                  55                  60  
 Ser Glu Ala Val Asn Gly Asp Asp Glu Glu Glu Glu Asp Gly Glu Glu  
 65                  70                  75                  80  
 Gln Gln His Lys Thr Lys Lys Arg Phe Lys Phe Thr Lys Met Ser Arg  
                   85                  90                  95  
 Lys Gln Thr Lys Lys Lys Val Pro Lys Val Ser Phe Ile Thr Arg Ser  
           100                  105                  110  
 Glu Val Leu His Leu Asp Asp Gly Tyr Lys Trp Arg Lys Tyr Gly Gln  
           115                  120                  125  
 Lys Pro Val Lys Asp Ser Pro Phe Pro Arg Asn Tyr Tyr Arg Cys Thr  
       130                  135                  140  
 Thr Thr Trp Cys Asp Val Lys Lys Arg Val Glu Arg Ser Phe Ser Asp  
 145                  150                  155                  160  
 Pro Ser Ser Val Ile Thr Thr Tyr Glu Gly Gln His Thr His Pro Arg  
                   165                  170                  175  
 Pro Leu Leu Ile Met Pro Lys Glu Gly Ser Ser Pro Ser Asn Gly Ser  
           180                  185                  190  
 Ala Ser Arg Ala His Ile Gly Leu Pro Thr Leu Pro Pro Gln Leu Leu  
       195                  200                  205  
 Asp Tyr Asn Asn Gln Gln Gln Gln Ala Pro Ser Ser Phe Gly Thr Glu  
       210                  215                  220  
 Tyr Ile Asn Arg Gln Glu Lys Gly Ile Asn His Asp Asp Asp Asp  
 225                  230                  235                  240  
 His Val Val Lys Lys Ser Arg Thr Arg Asp Leu Leu Asp Gly Ala Gly  
                   245                  250                  255  
 Leu Val Lys Asp His Gly Leu Leu Gln Asp Val Val Pro Ser His Ile  
           260                  265                  270  
 Ile Lys Glu Glu Tyr  
       275

&lt;210&gt; 145

&lt;211&gt; 795

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(795)

&lt;400&gt; 145

atg gac gtc tat ggc tta tct tca cca gac tta ctt cga atc gac gac 48  
 Met Asp Val Tyr Gly Leu Ser Ser Pro Asp Leu Leu Arg Ile Asp Asp  
   1                  5                  10                  15

ctt ctt gat ttc tcc aac gaa gac atc ttc tcc gct tct tct tcc ggt 96  
 Leu Leu Asp Phe Ser Asn Glu Asp Ile Phe Ser Ala Ser Ser Ser Gly  
           20                  25                  30

ggt tcc acc gcc gct act tcc tct tct tct ttc cct cct cct caa aac 144  
 Gly Ser Thr Ala Ala Thr Ser Ser Ser Ser Phe Pro Pro Pro Gln Asn  
       35                  40                  45



cct agt ttc cac cac cac cat ctc cct tcc tcc gcc gat cat cac tcc	192
Pro Ser Phe His His His His Leu Pro Ser Ser Ala Asp His His Ser	
50 55 60	
ttc ctc cac gac att tgc gtt ccc agt gat gac gca gct cat ctt gaa	240
Phe Leu His Asp Ile Cys Val Pro Ser Asp Asp Ala Ala His Leu Glu	
65 70 75 80	
tgg ctt tcg caa ttc gtg gac gat tct ttc gct gat ttt ccg gcg aat	288
Trp Leu Ser Gln Phe Val Asp Asp Ser Phe Ala Asp Phe Pro Ala Asn	
85 90 95	
cca tta gga gga act atg act tct gtc aaa act gaa act tcc ttt ccg	336
Pro Leu Gly Gly Thr Met Thr Ser Val Lys Thr Glu Thr Ser Phe Pro	
100 105 110	
ggg aaa cca aga agc aaa cga tca aga gct cct gct cct ttc gcc gga	384
Gly Lys Pro Arg Ser Lys Arg Ser Arg Ala Pro Ala Pro Phe Ala Gly	
115 120 125	
aca tgg tct ccg atg cca ctg gaa tcc gag cat cag cag ctt cac tcc	432
Thr Trp Ser Pro Met Pro Leu Glu Ser Glu His Gln Gln Leu His Ser	
130 135 140	
gcc gcc aaa ttc aag cca aag aaa gaa caa tcc ggc gga gga gga gga	480
Ala Ala Lys Phe Lys Pro Lys Lys Glu Gln Ser Gly Gly Gly Gly Gly	
145 150 155 160	
gga gga gga aga cat cag tca tcg tca tcg gag act acg gaa gga gga	528
Gly Gly Gly Arg His Gln Ser Ser Ser Ser Glu Thr Thr Glu Gly Gly	
165 170 175	
gga atg agg aga tgt act cac tgt gca tcg gag aaa acg cca cag tgg	576
Gly Met Arg Arg Cys Thr His Cys Ala Ser Glu Lys Thr Pro Gln Trp	
180 185 190	
agg aca gga cca ctt gga cct aaa aca cta tgt aac gct tgt gga gtc	624
Arg Thr Gly Pro Leu Gly Pro Lys Thr Leu Cys Asn Ala Cys Gly Val	
195 200 205	
cgg ttt aaa tcc ggt aga ctt gta ccg gaa tat aga ccg gct tcg agt	672
Arg Phe Lys Ser Gly Arg Leu Val Pro Glu Tyr Arg Pro Ala Ser Ser	
210 215 220	
cct act ttt gtt ttg act cag cat tca aac tct cac cgg aaa gtg atg	720
Pro Thr Phe Val Leu Thr Gln His Ser Asn Ser His Arg Lys Val Met	
225 230 235 240	
gag ctt cga cgg cag aaa gaa gtt atg aga caa cca caa caa gtt caa	768
Glu Leu Arg Arg Gln Lys Glu Val Met Arg Gln Pro Gln Gln Val Gln	
245 250 255	
ctt cat cac cac cac cac ccg ttt tag	795
Leu His His His His His Pro Phe *	
260	

<210> 146  
 <211> 264  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (178)...(214)  
 <223> Conserved domain

<400> 146  
 Met Asp Val Tyr Gly Leu Ser Ser Pro Asp Leu Leu Arg Ile Asp Asp  
 1 5 10 15  
 Leu Leu Asp Phe Ser Asn Glu Asp Ile Phe Ser Ala Ser Ser Ser Gly  
 20 25 30  
 Gly Ser Thr Ala Ala Thr Ser Ser Ser Phe Pro Pro Pro Gln Asn  
 35 40 45  
 Pro Ser Phe His His His His Leu Pro Ser Ser Ala Asp His His Ser  
 50 55 60  
 Phe Leu His Asp Ile Cys Val Pro Ser Asp Asp Ala Ala His Leu Glu  
 65 70 75 80  
 Trp Leu Ser Gln Phe Val Asp Asp Ser Phe Ala Asp Phe Pro Ala Asn  
 85 90 95  
 Pro Leu Gly Gly Thr Met Thr Ser Val Lys Thr Glu Thr Ser Phe Pro  
 100 105 110  
 Gly Lys Pro Arg Ser Lys Arg Ser Arg Ala Pro Ala Pro Phe Ala Gly  
 115 120 125  
 Thr Trp Ser Pro Met Pro Leu Glu Ser Glu His Gln Gln Leu His Ser  
 130 135 140  
 Ala Ala Lys Phe Lys Pro Lys Lys Glu Gln Ser Gly Gly Gly Gly Gly  
 145 150 155 160  
 Gly Gly Gly Arg His Gln Ser Ser Ser Ser Glu Thr Thr Glu Gly Gly  
 165 170 175  
 Gly Met Arg Arg Cys Thr His Cys Ala Ser Glu Lys Thr Pro Gln Trp  
 180 185 190  
 Arg Thr Gly Pro Leu Gly Pro Lys Thr Leu Cys Asn Ala Cys Gly Val  
 195 200 205  
 Arg Phe Lys Ser Gly Arg Leu Val Pro Glu Tyr Arg Pro Ala Ser Ser  
 210 215 220  
 Pro Thr Phe Val Leu Thr Gln His Ser Asn Ser His Arg Lys Val Met  
 225 230 235 240  
 Glu Leu Arg Arg Gln Lys Glu Val Met Arg Gln Pro Gln Gln Val Gln  
 245 250 255  
 Leu His His His His His Pro Phe  
 260

<210> 147  
 <211> 2026  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (297)...(1781)

&lt;400&gt; 147

caaaaaaaaaa gtttcaattt ttgaaagctc tgagaaatga aatctatcat tctctctctc 60  
 tatctctatc ttccttttca gatttcgctt cttcaattca tgaaatcctc gtgattctac 120  
 tttaatgctt ctcttttttt acttttccaa gtctctgaat attcaaagta tatatctttt 180  
 gttttcaaac ttttgacaaa ttgtcttcaa gcttccaaat ttcagttaaa ggtctcaact 240  
 ttgcagaatt ttcctctaaa gggtcagact ttgggggtaaa ggtgtcaact ttggcg atg 299

Met

1

ggt ctt gac gga aac aat ggt gga ggg gtt tgg tta aac ggt ggt ggt 347  
 Gly Leu Asp Gly Asn Asn Gly Gly Gly Val Trp Leu Asn Gly Gly Gly  
                   5                                  10                                  15

gga gaa agg gaa gag aac gag gaa ggt tca tgg gga agg aat caa gaa 395  
 Gly Glu Arg Glu Glu Asn Glu Glu Gly Ser Trp Gly Arg Asn Gln Glu  
                   20                                  25                                  30

gat ggt tct tct cag ttt aag cct atg ctt gaa ggt gat tgg ttt agt 443  
 Asp Gly Ser Ser Gln Phe Lys Pro Met Leu Glu Gly Asp Trp Phe Ser  
                   35                                  40                                  45

agt aac caa cca cat cca caa gat ctt cag atg tta cag aat cag cca 491  
 Ser Asn Gln Pro His Pro Gln Asp Leu Gln Met Leu Gln Asn Gln Pro  
                   50                                  55                                  60                                  65

gat ttc aga tac ttt ggt ggt ttt cct ttt aac cct aat gat aat ctt 539  
 Asp Phe Arg Tyr Phe Gly Gly Phe Pro Phe Asn Pro Asn Asp Asn Leu  
                                   70                                  75                                  80

ctt ctt caa cac tct att gat tct tct tct tct tgt tct cct tct caa 587  
 Leu Leu Gln His Ser Ile Asp Ser Ser Ser Ser Cys Ser Pro Ser Gln  
                                   85                                  90                                  95

gct ttt agt ctt gac cct tct cag caa aat cag ttc ttg tca act aac 635  
 Ala Phe Ser Leu Asp Pro Ser Gln Gln Asn Gln Phe Leu Ser Thr Asn  
                   100                                  105                                  110

aac aac aag ggt tgt ctt ctc aat gtt cct tct tct gca aac cct ttt 683  
 Asn Asn Lys Gly Cys Leu Leu Asn Val Pro Ser Ser Ala Asn Pro Phe  
                   115                                  120                                  125

gat aat gct ttt gag ttt ggc tct gaa tct ggt ttt ctt aac caa atc 731  
 Asp Asn Ala Phe Glu Phe Gly Ser Glu Ser Gly Phe Leu Asn Gln Ile  
                   130                                  135                                  140                                  145

cat gct cct att tcg atg ggg ttt ggt tct ttg aca caa ttg ggg aac 779  
 His Ala Pro Ile Ser Met Gly Phe Gly Ser Leu Thr Gln Leu Gly Asn  
                                   150                                  155                                  160

agg gat ttg agt tct gtt cct gat ttc ttg tct gct cgg tca ctt ctt 827  
 Arg Asp Leu Ser Ser Val Pro Asp Phe Leu Ser Ala Arg Ser Leu Leu  
                   165                                  170                                  175

gcg ccg gaa agc aac aac aac aac aca atg ttg tgt ggt ggt ttc aca 875  
 Ala Pro Glu Ser Asn Asn Asn Asn Thr Met Leu Cys Gly Gly Phe Thr  
                   180                                  185                                  190

gct ccg ttg gag ttg gaa ggt ttt ggt agt cct gct aat ggt ggt ttt	923
Ala Pro Leu Glu Leu Glu Gly Phe Gly Ser Pro Ala Asn Gly Gly Phe	
195 200 205	
ggt ggg aac aga gcg aaa gtt ctg aag cct tta gag gtg tta gca tcg	971
Val Gly Asn Arg Ala Lys Val Leu Lys Pro Leu Glu Val Leu Ala Ser	
210 215 220 225	
tct ggt gca cag cct act ctg ttc cag aaa cgt gca gct atg cgt cag	1019
Ser Gly Ala Gln Pro Thr Leu Phe Gln Lys Arg Ala Ala Met Arg Gln	
230 235 240	
agc tct gga agc aaa atg gga aat tcg gag agt tcg gga atg agg agg	1067
Ser Ser Gly Ser Lys Met Gly Asn Ser Glu Ser Ser Gly Met Arg Arg	
245 250 255	
ttt agt gat gat gga gat atg gat gag act ggg att gag gtt tct ggg	1115
Phe Ser Asp Asp Gly Asp Met Asp Glu Thr Gly Ile Glu Val Ser Gly	
260 265 270	
ttg aac tat gag tct gat gag ata aat gag agc ggt aaa gcg gct gag	1163
Leu Asn Tyr Glu Ser Asp Glu Ile Asn Glu Ser Gly Lys Ala Ala Glu	
275 280 285	
agt gtt cag att gga gga gga gga aag ggt aag aag aaa ggt atg cct	1211
Ser Val Gln Ile Gly Gly Gly Lys Gly Lys Lys Lys Gly Met Pro	
290 295 300 305	
gct aag aat ctg atg gct gag agg aga agg agg aag aag ctt aat gat	1259
Ala Lys Asn Leu Met Ala Glu Arg Arg Arg Arg Lys Lys Leu Asn Asp	
310 315 320	
agg ctt tat atg ctt aga tca gtt gtc ccc aag atc agc aaa atg gat	1307
Arg Leu Tyr Met Leu Arg Ser Val Val Pro Lys Ile Ser Lys Met Asp	
325 330 335	
aga gca tca ata ctt gga gat gca att gat tat ctg aag gaa ctt cta	1355
Arg Ala Ser Ile Leu Gly Asp Ala Ile Asp Tyr Leu Lys Glu Leu Leu	
340 345 350	
caa agg atc aat gat ctt cac aat gaa ctt gag tca act cct cct gga	1403
Gln Arg Ile Asn Asp Leu His Asn Glu Leu Glu Ser Thr Pro Pro Gly	
355 360 365	
tct ttg cct cca act tca tca agc ttc cat ccg ttg aca cct aca ccg	1451
Ser Leu Pro Pro Thr Ser Ser Phe His Pro Leu Thr Pro Thr Pro	
370 375 380 385	
caa act ctt tct tgt cgt gtc aag gaa gag ttg tgt ccc tct tct tta	1499
Gln Thr Leu Ser Cys Arg Val Lys Glu Glu Leu Cys Pro Ser Ser Leu	
390 395 400	
cca agt cct aaa ggc cag caa gct aga gtt gag gtt aga tta agg gaa	1547
Pro Ser Pro Lys Gly Gln Gln Ala Arg Val Glu Val Arg Leu Arg Glu	
405 410 415	
gga aga gca gtg aac att cat atg ttc tgt ggt cgt aga ccg ggt ctg	1595

Gly Arg Ala Val Asn Ile His Met Phe Cys Gly Arg Arg Pro Gly Leu  
 420 425 430

ttg ctc gct acc atg aaa gct ttg gat aat ctt gga ttg gat gtt cag 1643  
 Leu Leu Ala Thr Met Lys Ala Leu Asp Asn Leu Gly Leu Asp Val Gln  
 435 440 445

caa gct gtg atc agc tgt ttt aat ggg ttt gcc ttg gat gtt ttc cgc 1691  
 Gln Ala Val Ile Ser Cys Phe Asn Gly Phe Ala Leu Asp Val Phe Arg  
 450 455 460 465

gct gag caa tgc caa gaa gga caa gag ata ctg cct gat caa atc aaa 1739  
 Ala Glu Gln Cys Gln Glu Gly Gln Glu Ile Leu Pro Asp Gln Ile Lys  
 470 475 480

gca gtg ctt ttc gat aca gca ggg tat gct ggt atg atc tga 1781  
 Ala Val Leu Phe Asp Thr Ala Gly Tyr Ala Gly Met Ile \*  
 485 490

tctgatcctg acttcgagtc cattaagcat ctgttgaagc agagctagaa gaactaagtc 1841  
 ccttttaaate tgcaattttc ttctcaactt tttttcttat gtcataactt caatctaagc 1901  
 atgtaatgca attgcaaagc agagttgttt ttaaattaag cttttgagaa cttgaggttg 1961  
 ttgttgttgg atacataact tcaacctttt attagcaatg ttaacttcca tttatgtttc 2021  
 atctt 2026

&lt;210&gt; 148

&lt;211&gt; 494

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (308)...(359)

&lt;223&gt; Conserved domain

&lt;400&gt; 148

Met Gly Leu Asp Gly Asn Asn Gly Gly Gly Val Trp Leu Asn Gly Gly  
 1 5 10 15

Gly Gly Glu Arg Glu Glu Asn Glu Glu Gly Ser Trp Gly Arg Asn Gln  
 20 25 30

Glu Asp Gly Ser Ser Gln Phe Lys Pro Met Leu Glu Gly Asp Trp Phe  
 35 40 45

Ser Ser Asn Gln Pro His Pro Gln Asp Leu Gln Met Leu Gln Asn Gln  
 50 55 60

Pro Asp Phe Arg Tyr Phe Gly Gly Phe Pro Phe Asn Pro Asn Asp Asn  
 65 70 75 80

Leu Leu Leu Gln His Ser Ile Asp Ser Ser Ser Cys Ser Pro Ser  
 85 90 95

Gln Ala Phe Ser Leu Asp Pro Ser Gln Gln Asn Gln Phe Leu Ser Thr  
 100 105 110

Asn Asn Asn Lys Gly Cys Leu Leu Asn Val Pro Ser Ser Ala Asn Pro  
 115 120 125

Phe Asp Asn Ala Phe Glu Phe Gly Ser Glu Ser Gly Phe Leu Asn Gln  
 130 135 140

Ile His Ala Pro Ile Ser Met Gly Phe Gly Ser Leu Thr Gln Leu Gly  
 145 150 155 160

Asn Arg Asp Leu Ser Ser Val Pro Asp Phe Leu Ser Ala Arg Ser Leu

```

      165      170      175
Leu Ala Pro Glu Ser Asn Asn Asn Asn Thr Met Leu Cys Gly Gly Phe
      180      185      190
Thr Ala Pro Leu Glu Leu Glu Gly Phe Gly Ser Pro Ala Asn Gly Gly
      195      200      205
Phe Val Gly Asn Arg Ala Lys Val Leu Lys Pro Leu Glu Val Leu Ala
      210      215      220
Ser Ser Gly Ala Gln Pro Thr Leu Phe Gln Lys Arg Ala Ala Met Arg
      225      230      235      240
Gln Ser Ser Gly Ser Lys Met Gly Asn Ser Glu Ser Ser Gly Met Arg
      245      250      255
Arg Phe Ser Asp Asp Gly Asp Met Asp Glu Thr Gly Ile Glu Val Ser
      260      265      270
Gly Leu Asn Tyr Glu Ser Asp Glu Ile Asn Glu Ser Gly Lys Ala Ala
      275      280      285
Glu Ser Val Gln Ile Gly Gly Gly Gly Lys Gly Lys Lys Lys Gly Met
      290      295      300
Pro Ala Lys Asn Leu Met Ala Glu Arg Arg Arg Arg Lys Lys Leu Asn
      305      310      315      320
Asp Arg Leu Tyr Met Leu Arg Ser Val Val Pro Lys Ile Ser Lys Met
      325      330      335
Asp Arg Ala Ser Ile Leu Gly Asp Ala Ile Asp Tyr Leu Lys Glu Leu
      340      345      350
Leu Gln Arg Ile Asn Asp Leu His Asn Glu Leu Glu Ser Thr Pro Pro
      355      360      365
Gly Ser Leu Pro Pro Thr Ser Ser Ser Phe His Pro Leu Thr Pro Thr
      370      375      380
Pro Gln Thr Leu Ser Cys Arg Val Lys Glu Glu Leu Cys Pro Ser Ser
      385      390      395      400
Leu Pro Ser Pro Lys Gly Gln Gln Ala Arg Val Glu Val Arg Leu Arg
      405      410      415
Glu Gly Arg Ala Val Asn Ile His Met Phe Cys Gly Arg Arg Pro Gly
      420      425      430
Leu Leu Leu Ala Thr Met Lys Ala Leu Asp Asn Leu Gly Leu Asp Val
      435      440      445
Gln Gln Ala Val Ile Ser Cys Phe Asn Gly Phe Ala Leu Asp Val Phe
      450      455      460
Arg Ala Glu Gln Cys Gln Glu Gly Gln Glu Ile Leu Pro Asp Gln Ile
      465      470      475      480
Lys Ala Val Leu Phe Asp Thr Ala Gly Tyr Ala Gly Met Ile
      485      490

```

&lt;210&gt; 149

&lt;211&gt; 1146

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1146)

&lt;400&gt; 149

```

atg gat cta aca gac cgt cgt aac cct ttt aac aat ctt gtt ttt ccg 48
Met Asp Leu Thr Asp Arg Arg Asn Pro Phe Asn Asn Leu Val Phe Pro
  1             5             10             15

```

ccg ccg cct ccg ccg cca tcc acg acc ttc aca agc cct ata ttc cca	96
Pro Pro Pro Pro Pro Pro Ser Thr Thr Phe Thr Ser Pro Ile Phe Pro	
20 25 30	
cga aca agc tct tcc ggc acc aat ttc ccc att ctg gcc atc gca gtg	144
Arg Thr Ser Ser Ser Gly Thr Asn Phe Pro Ile Leu Ala Ile Ala Val	
35 40 45	
att gga atc tta gcc act gcg ttc tta ctt gta agt tac tac atc ttc	192
Ile Gly Ile Leu Ala Thr Ala Phe Leu Leu Val Ser Tyr Tyr Ile Phe	
50 55 60	
gtg atc aaa tgc tgt ctt aat tgg cac caa atc gac atc ttt cgc cgc	240
Val Ile Lys Cys Cys Leu Asn Trp His Gln Ile Asp Ile Phe Arg Arg	
65 70 75 80	
cgc aga cga agc agt gac caa aac cct cta atg att tac tct cct cat	288
Arg Arg Arg Ser Ser Asp Gln Asn Pro Leu Met Ile Tyr Ser Pro His	
85 90 95	
gag gta aac aga gga cta gac gaa tcc gcc att aga gct atc cca gtc	336
Glu Val Asn Arg Gly Leu Asp Glu Ser Ala Ile Arg Ala Ile Pro Val	
100 105 110	
ttc aaa ttc aag aag aga gac gtt gtt gca gga gaa gaa gat cag agt	384
Phe Lys Phe Lys Lys Arg Asp Val Val Ala Gly Glu Glu Asp Gln Ser	
115 120 125	
aag aac tct caa gaa tgc tct gtt tgt tta aac gag ttt caa gaa gac	432
Lys Asn Ser Gln Glu Cys Ser Val Cys Leu Asn Glu Phe Gln Glu Asp	
130 135 140	
gag aag cta agg att att cct aac tgc tgc cac gtg ttt cac att gat	480
Glu Lys Leu Arg Ile Ile Pro Asn Cys Cys His Val Phe His Ile Asp	
145 150 155 160	
tgc att gat atc tgg ctt cag ggc aac gca aat tgt ccc ttg tgc aga	528
Cys Ile Asp Ile Trp Leu Gln Gly Asn Ala Asn Cys Pro Leu Cys Arg	
165 170 175	
acc agc gtt tct tgc gaa gca agt ttc act ctt gac cta atc tct gca	576
Thr Ser Val Ser Cys Glu Ala Ser Phe Thr Leu Asp Leu Ile Ser Ala	
180 185 190	
ccg agc tct cct ccg gag aat agc cct cat tct cgg aac agg aat ctc	624
Pro Ser Ser Pro Arg Glu Asn Ser Pro His Ser Arg Asn Arg Asn Leu	
195 200 205	
gaa ccc ggc ctg gtt cta gga ggc gat gat gac ttc gtc gtc ata gag	672
Glu Pro Gly Leu Val Leu Gly Gly Asp Asp Asp Phe Val Val Ile Glu	
210 215 220	
ctt ggg gcc agt aat ggt aac aac aga gaa agc gtg aga aac ata gac	720
Leu Gly Ala Ser Asn Gly Asn Asn Arg Glu Ser Val Arg Asn Ile Asp	
225 230 235 240	
ttc ctt acg gag caa gaa agg gtt acc tcg aat gag gtc tcg acc gga	768

```
<210> 150
<211> 381
<212> PRT
<213> Arabidopsis thaliana
```

```
<220>
<221> DOMAIN
<222> (134) ... (159)
<223> Conserved domain
```

<400> 150															
Met	Asp	Leu	Thr	Asp	Arg	Arg	Asn	Pro	Phe	Asn	Asn	Leu	Val	Phe	Pro
1				5					10					15	
Pro	Pro	Pro	Pro	Pro	Pro	Ser	Thr	Thr	Phe	Thr	Ser	Pro	Ile	Phe	Pro
				20				25					30		
Arg	Thr	Ser	Ser	Ser	Gly	Thr	Asn	Phe	Pro	Ile	Leu	Ala	Ile	Ala	Val
		35					40					45			
Ile	Gly	Ile	Leu	Ala	Thr	Ala	Phe	Leu	Leu	Val	Ser	Tyr	Tyr	Ile	Phe
	50					55					60				
Val	Ile	Lys	Cys	Cys	Leu	Asn	Trp	His	Gln	Ile	Asp	Ile	Phe	Arg	Arg



```

65          70          75          80
Arg Arg Arg Ser Ser Asp Gln Asn Pro Leu Met Ile Tyr Ser Pro His
          85          90          95
Glu Val Asn Arg Gly Leu Asp Glu Ser Ala Ile Arg Ala Ile Pro Val
          100          105          110
Phe Lys Phe Lys Lys Arg Asp Val Ala Gly Glu Glu Asp Gln Ser
          115          120          125
Lys Asn Ser Gln Glu Cys Ser Val Cys Leu Asn Glu Phe Gln Glu Asp
          130          135          140
Glu Lys Leu Arg Ile Ile Pro Asn Cys Cys His Val Phe His Ile Asp
145          150          155          160
Cys Ile Asp Ile Trp Leu Gln Gly Asn Ala Asn Cys Pro Leu Cys Arg
          165          170          175
Thr Ser Val Ser Cys Glu Ala Ser Phe Thr Leu Asp Leu Ile Ser Ala
          180          185          190
Pro Ser Ser Pro Arg Glu Asn Ser Pro His Ser Arg Asn Arg Asn Leu
          195          200          205
Glu Pro Gly Leu Val Leu Gly Gly Asp Asp Asp Phe Val Val Ile Glu
          210          215          220
Leu Gly Ala Ser Asn Gly Asn Asn Arg Glu Ser Val Arg Asn Ile Asp
225          230          235          240
Phe Leu Thr Glu Gln Glu Arg Val Thr Ser Asn Glu Val Ser Thr Gly
          245          250          255
Asn Ser Pro Lys Ser Val Ser Pro Leu Pro Ile Lys Phe Gly Asn Arg
          260          265          270
Gly Met Tyr Lys Lys Glu Arg Lys Phe His Lys Val Thr Ser Met Gly
          275          280          285
Asp Glu Cys Ile Asp Thr Arg Gly Lys Asp Gly His Phe Gly Glu Ile
          290          295          300
Gln Pro Ile Arg Arg Ser Ile Ser Met Asp Ser Ser Val Asp Arg Gln
305          310          315          320
Leu Tyr Leu Ala Val Gln Glu Glu Ile Ser Arg Arg Asn Arg Gln Ile
          325          330          335
Pro Val Ala Gly Asp Gly Glu Asp Ser Ser Ser Ser Gly Gly Asn
          340          345          350
Ser Arg Val Met Lys Arg Cys Phe Phe Ser Phe Gly Ser Ser Arg Thr
          355          360          365
Ser Lys Ser Ser Ser Ile Leu Pro Val Tyr Leu Glu Pro
          370          375          380

```

<210> 151  
 <211> 1064  
 <212> DNA  
 <213> *Arabidopsis thaliana*

<220>  
 <221> CDS  
 <222> (61)...(849)

```

<400> 151
taaagaaaga gaaaaaaagc tttcgtagt tctattgaaa ccagagaaaa gccaaagggg 60
atg caa cca aca tcc gtc ggt agt agc ggc ggt ggt gac gac gga gga 108
Met Gln Pro Thr Ser Val Gly Ser Ser Gly Gly Gly Asp Asp Gly Gly
  1          5          10          15

ggc aga gga gga gga gga ggg cta agt aga agt gga cta tct cgg atc 156

```

Gly Arg Gly Gly Gly Gly Gly Leu Ser Arg Ser Gly Leu Ser Arg Ile	
20 25 30	
cgt tca gct cca gcg act tgg ctt gaa gct tta ctt gag gaa gat gaa	204
Arg Ser Ala Pro Ala Thr Trp Leu Glu Ala Leu Leu Glu Glu Asp Glu	
35 40 45	
gaa gag tct ttg aaa cct aat ctt ggt ctc acc gat ttg ctt acc ggg	252
Glu Glu Ser Leu Lys Pro Asn Leu Gly Leu Thr Asp Leu Leu Thr Gly	
50 55 60	
aac tcg aac gat tta ccg aca agt cgc ggc tcg ttc gag ttc ccg att	300
Asn Ser Asn Asp Leu Pro Thr Ser Arg Gly Ser Phe Glu Phe Pro Ile	
65 70 75 80	
cct gtt gag caa ggg ttg tat caa caa ggt ggg ttt cac cga cag aat	348
Pro Val Glu Gln Gly Leu Tyr Gln Gln Gly Gly Phe His Arg Gln Asn	
85 90 95	
agt act ccg gcg gat ttt ctt agt ggt tct gat gga ttt atc caa agc	396
Ser Thr Pro Ala Asp Phe Leu Ser Gly Ser Asp Gly Phe Ile Gln Ser	
100 105 110	
ttt ggg att cag gcg aat tac gat tac tta tcg ggg aat atc gat gtt	444
Phe Gly Ile Gln Ala Asn Tyr Asp Tyr Leu Ser Gly Asn Ile Asp Val	
115 120 125	
tct ccg gga agt aag ccg tct aga gaa atg gaa gca ctc ttc tct tct	492
Ser Pro Gly Ser Lys Arg Ser Arg Glu Met Glu Ala Leu Phe Ser Ser	
130 135 140	
cct gag ttt act tct caa atg aaa gga gag caa agc agc ggt caa gtt	540
Pro Glu Phe Thr Ser Gln Met Lys Gly Glu Gln Ser Ser Gly Gln Val	
145 150 155 160	
cct acc gga gta tca agc atg tcg gat atg aac atg gag aac ctt atg	588
Pro Thr Gly Val Ser Ser Met Ser Asp Met Asn Met Glu Asn Leu Met	
165 170 175	
gag gac tct gtt gct ttt agg gtt cgg gct aaa cgt ggt tgc gca act	636
Glu Asp Ser Val Ala Phe Arg Val Arg Ala Lys Arg Gly Cys Ala Thr	
180 185 190	
cat ccc cgc agc att gcc gag agg gta cga agg acg cgg att agt gat	684
His Pro Arg Ser Ile Ala Glu Arg Val Arg Arg Thr Arg Ile Ser Asp	
195 200 205	
cgg ata agg aag cta caa gag ctt gta cct aac atg gac aag caa acc	732
Arg Ile Arg Lys Leu Gln Glu Leu Val Pro Asn Met Asp Lys Gln Thr	
210 215 220	
aac act gca gac atg tta gaa gaa gca gta gaa tac gtg aaa gtt ctt	780
Asn Thr Ala Asp Met Leu Glu Glu Ala Val Glu Tyr Val Lys Val Leu	
225 230 235 240	
caa agg cag atc cag gag tta aca gaa gaa cag aag agg tgc aca tgc	828
Gln Arg Gln Ile Gln Glu Leu Thr Glu Glu Gln Lys Arg Cys Thr Cys	

245 250 255

ata cct aag gaa gaa caa taa ggtttgctcc tgatttgctt tatatttgct 879  
 Ile Pro Lys Glu Glu Gln \*  
 260

taacggcaat gatctgatcg aaaaattcga aagatgatct tagcttgaat ttagatggat 939  
 gtcattgtga aaagtatat atttgataaa tggatgtagg tgtaatatata aatttttgta 999  
 caataatgaa gaaagttaaa aagaattaat gaaaacatat attctttatg atataaaaaa 1059  
 aaaaa 1064

<210> 152  
 <211> 262  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (198)...(247)  
 <223> Conserved domain

<400> 152  
 Met Gln Pro Thr Ser Val Gly Ser Ser Gly Gly Gly Asp Asp Gly Gly  
 1 5 10 15  
 Gly Arg Gly Gly Gly Gly Gly Leu Ser Arg Ser Gly Leu Ser Arg Ile  
 20 25 30  
 Arg Ser Ala Pro Ala Thr Trp Leu Glu Ala Leu Leu Glu Glu Asp Glu  
 35 40 45  
 Glu Glu Ser Leu Lys Pro Asn Leu Gly Leu Thr Asp Leu Leu Thr Gly  
 50 55 60  
 Asn Ser Asn Asp Leu Pro Thr Ser Arg Gly Ser Phe Glu Phe Pro Ile  
 65 70 75 80  
 Pro Val Glu Gln Gly Leu Tyr Gln Gln Gly Gly Phe His Arg Gln Asn  
 85 90 95  
 Ser Thr Pro Ala Asp Phe Leu Ser Gly Ser Asp Gly Phe Ile Gln Ser  
 100 105 110  
 Phe Gly Ile Gln Ala Asn Tyr Asp Tyr Leu Ser Gly Asn Ile Asp Val  
 115 120 125  
 Ser Pro Gly Ser Lys Arg Ser Arg Glu Met Glu Ala Leu Phe Ser Ser  
 130 135 140  
 Pro Glu Phe Thr Ser Gln Met Lys Gly Glu Gln Ser Ser Gly Gln Val  
 145 150 155 160  
 Pro Thr Gly Val Ser Ser Met Ser Asp Met Asn Met Glu Asn Leu Met  
 165 170 175  
 Glu Asp Ser Val Ala Phe Arg Val Arg Ala Lys Arg Gly Cys Ala Thr  
 180 185 190  
 His Pro Arg Ser Ile Ala Glu Arg Val Arg Arg Thr Arg Ile Ser Asp  
 195 200 205  
 Arg Ile Arg Lys Leu Gln Glu Leu Val Pro Asn Met Asp Lys Gln Thr  
 210 215 220  
 Asn Thr Ala Asp Met Leu Glu Glu Ala Val Glu Tyr Val Lys Val Leu  
 225 230 235 240  
 Gln Arg Gln Ile Gln Glu Leu Thr Glu Glu Gln Lys Arg Cys Thr Cys  
 245 250 255  
 Ile Pro Lys Glu Glu Gln  
 260

<210> 153  
 <211> 831  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (201)...(791)

<400> 153  
 gtgtcgggtga attttttgaa acttcttctc ttttgcgggtt tcgtgttcca ctctctcttt 60  
 cttggcccac gtgttcatca atctctccct ccgcattgtaa tcgcttcgcc gtcaatatca 120  
 catctttctt cttctttatc tttaaaatct ctttagatcg attcttttgt ggattcttga 180  
 aatctccgga gaaaaccact atg gag acg gcg act gaa gtg gcc acg gtg gtg 233  
 Met Glu Thr Ala Thr Glu Val Ala Thr Val Val  
 1 5 10  
 tca act ccg gcg gtt acg gtt gcg gcg gtg gcg acg agg aag aga gat 281  
 Ser Thr Pro Ala Val Thr Val Ala Val Ala Thr Arg Lys Arg Asp  
 15 20 25  
 aag ccg tat aaa ggg ata agg atg agg aag tgg ggg aag tgg gtg gcg 329  
 Lys Pro Tyr Lys Gly Ile Arg Met Arg Lys Trp Gly Lys Trp Val Ala  
 30 35 40  
 gag ata aga gag cct aat aaa agg tca agg atc tgg ctt ggc tct tac 377  
 Glu Ile Arg Glu Pro Asn Lys Arg Ser Arg Ile Trp Leu Gly Ser Tyr  
 45 50 55  
 tct act cct gaa gcg gcg gcg cgt gct tac gac acg gcg gtg ttt tat 425  
 Ser Thr Pro Glu Ala Ala Ala Arg Ala Tyr Asp Thr Ala Val Phe Tyr  
 60 65 70 75  
 ctc cga ggt cct tct gct cgg ctt aac ttc ccg gag ctt tta gcc gga 473  
 Leu Arg Gly Pro Ser Ala Arg Leu Asn Phe Pro Glu Leu Leu Ala Gly  
 80 85 90  
 gtg acg gtg acg gga gga ggc gga gga gga gtg aac ggt ggt gga gat 521  
 Val Thr Val Thr Gly Gly Gly Gly Gly Val Asn Gly Gly Gly Asp  
 95 100 105  
 atg tcg gcg gcg tat ata agg aga aaa gcg gcg gag gtt gga gca caa 569  
 Met Ser Ala Ala Tyr Ile Arg Arg Lys Ala Ala Glu Val Gly Ala Gln  
 110 115 120  
 gtg gat gcg tta gaa gcg gcg ggg gcg gga ggg aat cgt cat cat cat 617  
 Val Asp Ala Leu Glu Ala Ala Gly Ala Gly Gly Asn Arg His His His  
 125 130 135  
 cat cat caa cat caa cgt ggt aat cat gat tac gta gat aat cat agt 665  
 His His Gln His Gln Arg Gly Asn His Asp Tyr Val Asp Asn His Ser  
 140 145 150 155  
 gat tat cgt att aat gat gat ctt atg gag tgt agt agt aaa gaa ggg 713  
 Asp Tyr Arg Ile Asn Asp Asp Leu Met Glu Cys Ser Ser Lys Glu Gly  
 160 165 170

ttt aag agg tgt aat gga tgc ttg gaa cgg gtt gat tta aac aaa tta 761  
Phe Lys Arg Cys Asn Gly Ser Leu Glu Arg Val Asp Leu Asn Lys Leu  
175 180 185

```

ccc gat ccg gaa act tca gat gac gat tag gaaagcaaaa aatagaaaca      811
Pro Asp Pro Glu Thr Ser Asp Asp Asp *
      190              195

```

aaaaaaaaaa aaaaaaaaaa 831

<210> 154

<211> 196

<212> PRT

<213> Arabidopsis thaliana

<400> 154

[illegible]

<210> 155

<211> 981

<212> DNA

<213> Arabidopsis thaliana

**<220>**

<221> CDS

<222> (36) ... (959)

<400> 155

gtaccggcga cctctttgtg ggtdactctt catca atg ggt gac aaa gga agg 53  
Met Gly Asp Lys Gly Arg

															1																5	
agc	tta	aag	atc	aac	aag	aac	atg	gag	gaa	ttc	acg	aaa	gtg	gaa	gaa	101																
Ser	Leu	Lys	Ile	Asn	Lys	Asn	Met	Glu	Glu	Phe	Thr	Lys	Val	Glu	Glu																	
				10					15					20																		
gaa	atg	gac	gta	agg	aga	ggg	cca	tgg	aca	gtt	gag	gaa	gat	tta	gag	149																
Glu	Met	Asp	Val	Arg	Arg	Gly	Pro	Trp	Thr	Val	Glu	Glu	Asp	Leu	Glu																	
				25					30					35																		
ctc	atc	aat	tac	att	gct	agt	cat	ggg	gaa	ggg	cga	tgg	aac	tct	ctc	197																
Leu	Ile	Asn	Tyr	Ile	Ala	Ser	His	Gly	Glu	Gly	Arg	Trp	Asn	Ser	Leu																	
				40					45					50																		
gct	cgt	tgc	gcc	gaa	ctc	aaa	agg	acc	gga	aaa	agc	tgc	aga	ctt	cgg	245																
Ala	Arg	Cys	Ala	Glu	Leu	Lys	Arg	Thr	Gly	Lys	Ser	Cys	Arg	Leu	Arg																	
				55					60					65																		
tgg	ctg	aac	tat	ctc	cga	cca	gat	gtg	cgc	cgt	gga	aac	ata	acc	ctc	293																
Trp	Leu	Asn	Tyr	Leu	Arg	Pro	Asp	Val	Arg	Arg	Gly	Asn	Ile	Thr	Leu																	
				75					80					85																		
gaa	gaa	caa	ctc	ttg	att	ctt	gaa	ctt	cac	aca	cgt	tgg	ggc	aat	aga	341																
Glu	Glu	Gln	Leu	Leu	Ile	Leu	Glu	Leu	His	Thr	Arg	Trp	Gly	Asn	Arg																	
				90					95					100																		
tgg	tct	aag	att	gca	caa	tat	tta	cca	gga	aga	acg	gat	aac	gag	atc	389																
Trp	Ser	Lys	Ile	Ala	Gln	Tyr	Leu	Pro	Gly	Arg	Thr	Asp	Asn	Glu	Ile																	
				105					110					115																		
aaa	aac	tat	tgg	aga	aca	cgt	gtt	caa	aag	cat	gca	aaa	cag	ctt	aaa	437																
Lys	Asn	Tyr	Trp	Arg	Thr	Arg	Val	Gln	Lys	His	Ala	Lys	Gln	Leu	Lys																	
				120					125					130																		
tgc	gac	gtg	aac	agt	caa	caa	ttt	aaa	gac	acc	atg	aag	tat	ctt	tgg	485																
Cys	Asp	Val	Asn	Ser	Gln	Gln	Phe	Lys	Asp	Thr	Met	Lys	Tyr	Leu	Trp																	
				135					140					145																		
atg	cct	cgg	ctc	gta	gaa	agg	atc	caa	gcc	gcg	tcc	atc	ggg	tct	gtt	533																
Met	Pro	Arg	Leu	Val	Glu	Arg	Ile	Gln	Ala	Ala	Ser	Ile	Gly	Ser	Val																	
				155					160					165																		
tcc	atg	tca	tct	tgc	gtc	acc	acc	tcc	tca	gat	cag	ttc	gtg	atc	aac	581																
Ser	Met	Ser	Ser	Cys	Val	Thr	Thr	Ser	Ser	Asp	Gln	Phe	Val	Ile	Asn																	
				170					175					180																		
aac	aac	aac	acc	aac	aac	gtg	gat	aat	ttg	gct	tta	atg	agt	aac	cct	629																
Asn	Asn	Asn	Thr	Asn	Asn	Val	Asp	Asn	Leu	Ala	Leu	Met	Ser	Asn	Pro																	
				185					190					195																		
aat	ggg	tac	atc	acg	ccg	gat	aat	tcc	agc	gtg	gca	gta	tct	cct	gta	677																
Asn	Gly	Tyr	Ile	Thr	Pro	Asp	Asn	Ser	Ser	Val	Ala	Val	Ser	Pro	Val																	
				200					205					210																		
tca	gat	ttg	acg	gag	tgt	caa	gtg	agt	agt	gaa	gtg	tgg	aag	att	ggg	725																
Ser	Asp	Leu	Thr	Glu	Cys	Gln	Val	Ser	Ser	Glu	Val	Trp	Lys	Ile	Gly																	
				215					220					225																		
												230																				

cag gat gag aat ttg gtg gat cca aaa atg aca tcg ccg aat tat atg 773  
 Gln Asp Glu Asn Leu Val Asp Pro Lys Met Thr Ser Pro Asn Tyr Met  
 235 240 245

gat aat agc agt gga cta tta aac gga gat ttt acg aag atg caa gat 821  
 Asp Asn Ser Ser Gly Leu Leu Asn Gly Asp Phe Thr Lys Met Gln Asp  
 250 255 260

caa agt gac ctt aat tgg ttt gaa aat att aat ggg atg gta cca aat 869  
 Gln Ser Asp Leu Asn Trp Phe Glu Asn Ile Asn Gly Met Val Pro Asn  
 265 270 275

tat tcg gac agt ttt tgg aac att gga aat gat gaa gac ttc tgg ctc 917  
 Tyr Ser Asp Ser Phe Trp Asn Ile Gly Asn Asp Glu Asp Phe Trp Leu  
 280 285 290

tta caa caa cat caa caa gtc cac gac aat gga agc ttc tga 959  
 Leu Gln Gln His Gln Val His Asp Asn Gly Ser Phe \*  
 295 300 305

atagacaaga agctatgcgg cc 981

<210> 156  
 <211> 307  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (28)...(134)  
 <223> Conserved domain

<400> 156  
 Met Gly Asp Lys Gly Arg Ser Leu Lys Ile Asn Lys Asn Met Glu Glu  
 1 5 10 15  
 Phe Thr Lys Val Glu Glu Glu Met Asp Val Arg Arg Gly Pro Trp Thr  
 20 25 30  
 Val Glu Glu Asp Leu Glu Leu Ile Asn Tyr Ile Ala Ser His Gly Glu  
 35 40 45  
 Gly Arg Trp Asn Ser Leu Ala Arg Cys Ala Glu Leu Lys Arg Thr Gly  
 50 55 60  
 Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp Val Arg  
 65 70 75 80  
 Arg Gly Asn Ile Thr Leu Glu Glu Gln Leu Leu Ile Leu Glu Leu His  
 85 90 95  
 Thr Arg Trp Gly Asn Arg Trp Ser Lys Ile Ala Gln Tyr Leu Pro Gly  
 100 105 110  
 Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Arg Thr Arg Val Gln Lys  
 115 120 125  
 His Ala Lys Gln Leu Lys Cys Asp Val Asn Ser Gln Gln Phe Lys Asp  
 130 135 140  
 Thr Met Lys Tyr Leu Trp Met Pro Arg Leu Val Glu Arg Ile Gln Ala  
 145 150 155 160  
 Ala Ser Ile Gly Ser Val Ser Met Ser Ser Cys Val Thr Thr Ser Ser  
 165 170 175  
 Asp Gln Phe Val Ile Asn Asn Asn Asn Thr Asn Asn Val Asp Asn Leu

```

180      185      190
Ala Leu Met Ser Asn Pro Asn Gly Tyr Ile Thr Pro Asp Asn Ser Ser
195      200      205
Val Ala Val Ser Pro Val Ser Asp Leu Thr Glu Cys Gln Val Ser Ser
210      215      220
Glu Val Trp Lys Ile Gly Gln Asp Glu Asn Leu Val Asp Pro Lys Met
225      230      235      240
Thr Ser Pro Asn Tyr Met Asp Asn Ser Ser Gly Leu Leu Asn Gly Asp
245      250      255
Phe Thr Lys Met Gln Asp Gln Ser Asp Leu Asn Trp Phe Glu Asn Ile
260      265      270
Asn Gly Met Val Pro Asn Tyr Ser Asp Ser Phe Trp Asn Ile Gly Asn
275      280      285
Asp Glu Asp Phe Trp Leu Leu Gln Gln His Gln Gln Val His Asp Asn
290      295      300
Gly Ser Phe
305

```

```

<210> 157
<211> 1891
<212> DNA
<213> Arabidopsis thaliana

```

```

<220>
<221> CDS
<222> (108)...(1664)

```

```

<400> 157
catcatgagc gtatgagagt aaattcttct cacatattaa taacaaccct tcaaagttat 60
aagattttta gagagagagc taccacgttt tcgtatctcc gggaacg atg gat gaa 116
                                     Met Asp Glu
                                     1

tca agt att att ccg gca gag aaa gtg gcc gga gct gag aaa aaa gag 164
Ser Ser Ile Ile Pro Ala Glu Lys Val Ala Gly Ala Glu Lys Lys Glu
5      10      15

ctt caa ggg ctg ctt aag acg gcg gtt caa tct gtg gac tgg act tat 212
Leu Gln Gly Leu Leu Lys Thr Ala Val Gln Ser Val Asp Trp Thr Tyr
20      25      30      35

agt gtc ttc tgg caa ttt tgt cct caa caa cgg gtc ttg gtg tgg ggg 260
Ser Val Phe Trp Gln Phe Cys Pro Gln Gln Arg Val Leu Val Trp Gly
40      45      50

aat gga tac tac aac ggt gca ata aag acg agg aag aca act caa cca 308
Asn Gly Tyr Tyr Asn Gly Ala Ile Lys Thr Arg Lys Thr Thr Gln Pro
55      60      65

gcg gag gtg acg gca gaa gag gcg gcg tta gag agg agc caa cag ctc 356
Ala Glu Val Thr Ala Glu Glu Ala Ala Leu Glu Arg Ser Gln Gln Leu
70      75      80

agg gag ctt tat gag aca ctt tta gcc gga gag tca acg tca gaa gca 404
Arg Glu Leu Tyr Glu Thr Leu Leu Ala Gly Glu Ser Thr Ser Glu Ala
85      90      95

```



aga gca tgc acc gca ttg tca ccg gag gat ttg acg gag aca gaa tgg	452
Arg Ala Cys Thr Ala Leu Ser Pro Glu Asp Leu Thr Glu Thr Glu Trp	
100 105 110 115	
ttt tat cta atg tgc gtc tct ttc tct ttt cct cct cca tct ggg atg	500
Phe Tyr Leu Met Cys Val Ser Phe Ser Phe Pro Pro Pro Ser Gly Met	
120 125 130	
cca gga aaa gcg tat gca agg agg aag cac gta tgg cta agt ggt gca	548
Pro Gly Lys Ala Tyr Ala Arg Arg Lys His Val Trp Leu Ser Gly Ala	
135 140 145	
aat gaà gtt gac agt aaa act ttt tct aga gct att ctc gct aag agt	596
Asn Glu Val Asp Ser Lys Thr Phe Ser Arg Ala Ile Leu Ala Lys Ser	
150 155 160	
gct aaa att cag aca gtg gtt tgc att cca atg ctt gat ggt gtt gtg	644
Ala Lys Ile Gln Thr Val Val Cys Ile Pro Met Leu Asp Gly Val Val	
165 170 175	
gaa cta ggc aca acg aaa aag gta aga gaa gat gta gag ttt gtt gag	692
Glu Leu Gly Thr Thr Lys Lys Val Arg Glu Asp Val Glu Phe Val Glu	
180 185 190 195	
ctc aca aag agt ttc ttc tat gac cac tgc aag acg aac cca aag ccg	740
Leu Thr Lys Ser Phe Phe Tyr Asp His Cys Lys Thr Asn Pro Lys Pro	
200 205 210	
gct ctt tct gaa cac tcc acc tac gaa gtg cat gaa gaa gcc gaa gac	788
Ala Leu Ser Glu His Ser Thr Tyr Glu Val His Glu Glu Ala Glu Asp	
215 220 225	
gaa gaa gaa gta gaa gaa gag atg aca atg tca gag gaa atg agg ctt	836
Glu Glu Glu Val Glu Glu Glu Met Thr Met Ser Glu Glu Met Arg Leu	
230 235 240	
ggc tct cct gat gat gaa gat gtt tcc aat caa aat cta cac tct gat	884
Gly Ser Pro Asp Asp Glu Asp Val Ser Asn Gln Asn Leu His Ser Asp	
245 250 255	
ctt cat att gaa tca acc cat acg tta gac aca cat atg gac atg atg	932
Leu His Ile Glu Ser Thr His Thr Leu Asp Thr His Met Asp Met Met	
260 265 270 275	
aat cta atg gag gaa ggt gga aac tat tct cag aca gta aca aca ctt	980
Asn Leu Met Glu Glu Gly Gly Asn Tyr Ser Gln Thr Val Thr Thr Leu	
280 285 290	
ctc atg tca cac ccc aca agt ctt ctt tca gat tca gtt tcc aca tct	1028
Leu Met Ser His Pro Thr Ser Leu Leu Ser Asp Ser Val Ser Thr Ser	
295 300 305	
tct tac atc caa tca tcg ttt gcc acg tgg agg gtt gag aat ggc aaa	1076
Ser Tyr Ile Gln Ser Ser Phe Ala Thr Trp Arg Val Glu Asn Gly Lys	
310 315 320	

gag cat cag caa gtg aaa acg gcg ccg tcg tca caa tgg gtg ctc aaa 1124  
 Glu His Gln Gln Val Lys Thr Ala Pro Ser Ser Gln Trp Val Leu Lys  
 325 330 335

caa atg atc ttc aga gtt cct ttc ctc cat gac aac act aaa gat aag 1172  
 Gln Met Ile Phe Arg Val Pro Phe Leu His Asp Asn Thr Lys Asp Lys  
 340 345 350 355

agg cta ccg cgg gaa gat ctg agc cac gta gta gca gag cga cgc agg 1220  
 Arg Leu Pro Arg Glu Asp Leu Ser His Val Val Ala Glu Arg Arg Arg  
 360 365 370

agg gag aag ctg aac gag aaa ttc ata acg ttg aga tca atg gtt cca 1268  
 Arg Glu Lys Leu Asn Glu Lys Phe Ile Thr Leu Arg Ser Met Val Pro  
 375 380 385

ttt gtg acc aag atg gat aaa gtc tca atc ctt gga gac acc att gcg 1316  
 Phe Val Thr Lys Met Asp Lys Val Ser Ile Leu Gly Asp Thr Ile Ala  
 390 395 400

tac gta aat cat ctt cga aag agg gtc cat gag ctt gag aat act cat 1364  
 Tyr Val Asn His Leu Arg Lys Arg Val His Glu Leu Glu Asn Thr His  
 405 410 415

cat gag caa cag cat aag cgg acg cgt act tgt aag aga aaa aca tcg 1412  
 His Glu Gln Gln His Lys Arg Thr Arg Thr Cys Lys Arg Lys Thr Ser  
 420 425 430 435

gag gag gtg gag gtt tcc atc ata gag aat gat gtt ttg tta gag atg 1460  
 Glu Glu Val Glu Val Ser Ile Ile Glu Asn Asp Val Leu Leu Glu Met  
 440 445 450

aga tgt gag tac cga gat ggt ttg ttg ctt gac att ctt cag gtt ctt 1508  
 Arg Cys Glu Tyr Arg Asp Gly Leu Leu Leu Asp Ile Leu Gln Val Leu  
 455 460 465

cat gag ctt ggt ata gag act acg gca gtt cat acc tcg gtg aac gac 1556  
 His Glu Leu Gly Ile Glu Thr Thr Ala Val His Thr Ser Val Asn Asp  
 470 475 480

cat gat ttc gag gcg gag ata agg gcg aaa gta aga ggg aag aaa gca 1604  
 His Asp Phe Glu Ala Glu Ile Arg Ala Lys Val Arg Gly Lys Lys Ala  
 485 490 495

agc atc gct gag gtc aaa aga gcc atc cac caa gtc ata ata cat gat 1652  
 Ser Ile Ala Glu Val Lys Arg Ala Ile His Gln Val Ile Ile His Asp  
 500 505 510 515

act aat cta tag accctaactt tattgatgcc aactctagag aaggataatt 1704  
 Thr Asn Leu \*

aagcgtatatt ttgttttagc ctcacatgta ttaagacatc agttacatat atagccggat 1764  
 gcaacatata aatgaaaatg tactagatga tattgttcat ttgtccaatg tagtacttgt 1824  
 gtatgatgca attgcaacat ataaatgcaa atgtactaga tgacgatgtt gttcgttgct 1884  
 caattta 1891

<210> 158  
 <211> 518  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (353)...(420)  
 <223> Conserved domain

<400> 158

```

Met Asp Glu Ser Ser Ile Ile Pro Ala Glu Lys Val Ala Gly Ala Glu
 1           5           10           15
Lys Lys Glu Leu Gln Gly Leu Leu Lys Thr Ala Val Gln Ser Val Asp
           20           25           30
Trp Thr Tyr Ser Val Phe Trp Gln Phe Cys Pro Gln Gln Arg Val Leu
           35           40           45
Val Trp Gly Asn Gly Tyr Tyr Asn Gly Ala Ile Lys Thr Arg Lys Thr
           50           55           60
Thr Gln Pro Ala Glu Val Thr Ala Glu Glu Ala Leu Glu Arg Ser
           65           70           75           80
Gln Gln Leu Arg Glu Leu Tyr Glu Thr Leu Leu Ala Gly Glu Ser Thr
           85           90           95
Ser Glu Ala Arg Ala Cys Thr Ala Leu Ser Pro Glu Asp Leu Thr Glu
           100          105          110
Thr Glu Trp Phe Tyr Leu Met Cys Val Ser Phe Ser Phe Pro Pro Pro
           115          120          125
Ser Gly Met Pro Gly Lys Ala Tyr Ala Arg Arg Lys His Val Trp Leu
           130          135          140
Ser Gly Ala Asn Glu Val Asp Ser Lys Thr Phe Ser Arg Ala Ile Leu
           145          150          155          160
Ala Lys Ser Ala Lys Ile Gln Thr Val Val Cys Ile Pro Met Leu Asp
           165          170          175
Gly Val Val Glu Leu Gly Thr Thr Lys Lys Val Arg Glu Asp Val Glu
           180          185          190
Phe Val Glu Leu Thr Lys Ser Phe Phe Tyr Asp His Cys Lys Thr Asn
           195          200          205
Pro Lys Pro Ala Leu Ser Glu His Ser Thr Tyr Glu Val His Glu Glu
           210          215          220
Ala Glu Asp Glu Glu Glu Val Glu Glu Glu Met Thr Met Ser Glu Glu
           225          230          235          240
Met Arg Leu Gly Ser Pro Asp Asp Glu Asp Val Ser Asn Gln Asn Leu
           245          250          255
His Ser Asp Leu His Ile Glu Ser Thr His Thr Leu Asp Thr His Met
           260          265          270
Asp Met Met Asn Leu Met Glu Glu Gly Gly Asn Tyr Ser Gln Thr Val
           275          280          285
Thr Thr Leu Leu Met Ser His Pro Thr Ser Leu Leu Ser Asp Ser Val
           290          295          300
Ser Thr Ser Ser Tyr Ile Gln Ser Ser Phe Ala Thr Trp Arg Val Glu
           305          310          315          320
Asn Gly Lys Glu His Gln Gln Val Lys Thr Ala Pro Ser Ser Gln Trp
           325          330          335
Val Leu Lys Gln Met Ile Phe Arg Val Pro Phe Leu His Asp Asn Thr
           340          345          350
Lys Asp Lys Arg Leu Pro Arg Glu Asp Leu Ser His Val Val Ala Glu
           355          360          365

```

Arg Arg Arg Arg Glu Lys Leu Asn Glu Lys Phe Ile Thr Leu Arg Ser  
 370 375 380  
 Met Val Pro Phe Val Thr Lys Met Asp Lys Val Ser Ile Leu Gly Asp  
 385 390 395 400  
 Thr Ile Ala Tyr Val Asn His Leu Arg Lys Arg Val His Glu Leu Glu  
 405 410 415  
 Asn Thr His His Glu Gln Gln His Lys Arg Thr Arg Thr Cys Lys Arg  
 420 425 430  
 Lys Thr Ser Glu Glu Val Glu Val Ser Ile Ile Glu Asn Asp Val Leu  
 435 440 445  
 Leu Glu Met Arg Cys Glu Tyr Arg Asp Gly Leu Leu Leu Asp Ile Leu  
 450 455 460  
 Gln Val Leu His Glu Leu Gly Ile Glu Thr Thr Ala Val His Thr Ser  
 465 470 475 480  
 Val Asn Asp His Asp Phe Glu Ala Glu Ile Arg Ala Lys Val Arg Gly  
 485 490 495  
 Lys Lys Ala Ser Ile Ala Glu Val Lys Arg Ala Ile His Gln Val Ile  
 500 505 510  
 Ile His Asp Thr Asn Leu  
 515

&lt;210&gt; 159

&lt;211&gt; 558

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(558)

&lt;400&gt; 159

atg act cga ccg tca aga tta ctt gag acg gcg gcg cca cca cca caa 48  
 Met Thr Arg Pro Ser Arg Leu Leu Glu Thr Ala Ala Pro Pro Pro Gln  
 1 5 10 15  
 ccg tcg gag gag atg atc gca gcg gaa tcc gac atg gtg gtg atc ttg 96  
 Pro Ser Glu Glu Met Ile Ala Ala Glu Ser Asp Met Val Val Ile Leu  
 20 25 30  
 tcg gct ctt ctt tgc gct ctt atc tgc gtt gct ggt ctc gcc gcc gtc 144  
 Ser Ala Leu Leu Cys Ala Leu Ile Cys Val Ala Gly Leu Ala Ala Val  
 35 40 45  
 gta cga tgc gct tgg ctc cgg cgg ttt aca gcc gga gga gat tcg ccg 192  
 Val Arg Cys Ala Trp Leu Arg Arg Phe Thr Ala Gly Gly Asp Ser Pro  
 50 55 60  
 tca ccg aac aaa ggc ttg aaa aag aaa gct ctt cag tct ctt cca aga 240  
 Ser Pro Asn Lys Gly Leu Lys Lys Lys Ala Leu Gln Ser Leu Pro Arg  
 65 70 75 80  
 tcc act ttc acc gcc gcg gaa tca acc tcc gcc gcc gcc gct gaa gag 288  
 Ser Thr Phe Thr Ala Ala Glu Ser Thr Ser Gly Ala Ala Ala Glu Glu  
 85 90 95  
 gga gac tcg acg gaa tgt gct att tgc ctc act gac ttc gcc gac ggt 336

Gly Asp Ser Thr Glu Cys Ala Ile Cys Leu Thr Asp Phe Ala Asp Gly  
                   100                  105                  110  
 gaa gaa ata aga gtg ctt cct ctt tgt ggt cat tct ttc cac gtg gag 384  
 Glu Glu Ile Arg Val Leu Pro Leu Cys Gly His Ser Phe His Val Glu  
                   115                  120                  125  
 tgt att gac aaa tgg cta gtt tct agg tct tct tgt cct tct tgt cgc 432  
 Cys Ile Asp Lys Trp Leu Val Ser Arg Ser Ser Cys Pro Ser Cys Arg  
                   130                  135                  140  
 agg att ctt acg ccg gtg aga tgt gac cgg tgt ggt cat gct tct acg 480  
 Arg Ile Leu Thr Pro Val Arg Cys Asp Arg Cys Gly His Ala Ser Thr  
                   145                  150                  155                  160  
 gcg gag atg aaa gat caa gct cat cgt cat caa cat cac caa cac tct 528  
 Ala Glu Met Lys Asp Gln Ala His Arg His Gln His His Gln His Ser  
                   165                  170                  175  
 tct act acc att cct acg ttt ctt cct taa 558  
 Ser Thr Thr Ile Pro Thr Phe Leu Pro \*  
                   180                  185

&lt;210&gt; 160

&lt;211&gt; 185

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (114)...(155)

&lt;223&gt; Conserved domain

&lt;400&gt; 160

Met Thr Arg Pro Ser Arg Leu Leu Glu Thr Ala Ala Pro Pro Pro Gln  
   1                  5                  10                  15  
 Pro Ser Glu Glu Met Ile Ala Ala Glu Ser Asp Met Val Val Ile Leu  
                   20                  25                  30  
 Ser Ala Leu Leu Cys Ala Leu Ile Cys Val Ala Gly Leu Ala Ala Val  
                   35                  40                  45  
 Val Arg Cys Ala Trp Leu Arg Arg Phe Thr Ala Gly Gly Asp Ser Pro  
                   50                  55                  60  
 Ser Pro Asn Lys Gly Leu Lys Lys Lys Ala Leu Gln Ser Leu Pro Arg  
   65                  70                  75                  80  
 Ser Thr Phe Thr Ala Ala Glu Ser Thr Ser Gly Ala Ala Ala Glu Glu  
                   85                  90                  95  
 Gly Asp Ser Thr Glu Cys Ala Ile Cys Leu Thr Asp Phe Ala Asp Gly  
                   100                  105                  110  
 Glu Glu Ile Arg Val Leu Pro Leu Cys Gly His Ser Phe His Val Glu  
                   115                  120                  125  
 Cys Ile Asp Lys Trp Leu Val Ser Arg Ser Ser Cys Pro Ser Cys Arg  
                   130                  135                  140  
 Arg Ile Leu Thr Pro Val Arg Cys Asp Arg Cys Gly His Ala Ser Thr  
   145                  150                  155                  160  
 Ala Glu Met Lys Asp Gln Ala His Arg His Gln His His Gln His Ser  
                   165                  170                  175

Ser Thr Thr Ile Pro Thr Phe Leu Pro  
180 185

<210> 161

<211> 2001

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (175)...(1878)

<400> 161

tgcttaattc caatgccatc gtgacgatt catctctctc tctctcttcc aattttccca 60  
attctttttt aaaaccctaa tttttcagat atctgattat ctcttgatt tcttctactc 120  
gatttgctcc cataaaaacc cttactttct tcaagttctg gttttcaccg attg atg 177  
Met  
1

ggt cgt ggc tca gtg acg tcg ctt gct cct ggg ttc cgt ttt cac ccg 225  
Gly Arg Gly Ser Val Thr Ser Leu Ala Pro Gly Phe Arg Phe His Pro  
5 10 15

acg gat gag gaa ctt gtt cgc tac tac ctt aag cgt aag gtc tgc aac 273  
Thr Asp Glu Glu Leu Val Arg Tyr Tyr Leu Lys Arg Lys Val Cys Asn  
20 25 30

aaa ccc ttt aag ttc gat gct att tcc gtc acc gac ata tac aag tct 321  
Lys Pro Phe Lys Phe Asp Ala Ile Ser Val Thr Asp Ile Tyr Lys Ser  
35 40 45

gag cct tgg gat cta cca gat aag tcg aag ctg aaa agt aga gac ttg 369  
Glu Pro Trp Asp Leu Pro Asp Lys Ser Lys Leu Lys Ser Arg Asp Leu  
50 55 60 65

gaa tgg tac ttc ttt agt atg ctg gat aag aag tac agt aat ggt tcc 417  
Glu Trp Tyr Phe Phe Ser Met Leu Asp Lys Lys Tyr Ser Asn Gly Ser  
70 75 80

aag acg aat cgt gct acg gag aaa ggg tat tgg aag acg act ggg aaa 465  
Lys Thr Asn Arg Ala Thr Glu Lys Gly Tyr Trp Lys Thr Thr Gly Lys  
85 90 95

gat cgg gag att cgt aat ggt tca aga gtc gtt ggg atg aag aag aca 513  
Asp Arg Glu Ile Arg Asn Gly Ser Arg Val Val Gly Met Lys Lys Thr  
100 105 110

ctt gtt tat cac aag ggt cga gct cct cgt ggt gaa agg acc aat tgg 561  
Leu Val Tyr His Lys Gly Arg Ala Pro Arg Gly Glu Arg Thr Asn Trp  
115 120 125

gtt atg cat gag tat cgg ctt tct gat gag gac ttg aag aaa gct ggt 609  
Val Met His Glu Tyr Arg Leu Ser Asp Glu Asp Leu Lys Lys Ala Gly  
130 135 140 145

gtg cca caa gaa gca tat gtg tta tgt agg ata ttc cag aaa agt ggt 657

Val	Pro	Gln	Glu	Ala	Tyr	Val	Leu	Cys	Arg	Ile	Phe	Gln	Lys	Ser	Gly		
				150					155						160		
acg	ggt	cct	aag	aat	ggg	gag	cag	tat	ggt	gct	cct	tat	ctt	gag	gag	705	
Thr	Gly	Pro	Lys	Asn	Gly	Glu	Gln	Tyr	Gly	Ala	Pro	Tyr	Leu	Glu	Glu		
			165					170					175				
gag	tgg	gaa	gaa	gat	gga	atg	act	tat	gta	cct	gct	caa	gat	gct	ttc	753	
Glu	Trp	Glu	Glu	Asp	Gly	Met	Thr	Tyr	Val	Pro	Ala	Gln	Asp	Ala	Phe		
		180					185					190					
agt	gaa	gga	ttg	gct	ttg	aat	gat	gat	gtt	tat	gtc	gat	att	gat	gac	801	
Ser	Glu	Gly	Leu	Ala	Leu	Asn	Asp	Asp	Val	Tyr	Val	Asp	Ile	Asp	Asp		
	195					200					205						
att	gac	gag	aag	ccc	gaa	aat	ctg	gtg	gtc	tat	gat	gcc	gtt	cct	att	849	
Ile	Asp	Glu	Lys	Pro	Glu	Asn	Leu	Val	Val	Tyr	Asp	Ala	Val	Pro	Ile		
210					215				220					225			
cta	cct	aac	tat	tgt	cat	ggg	gaa	tca	agt	aac	aat	gtt	gaa	tca	ggc	897	
Leu	Pro	Asn	Tyr	Cys	His	Gly	Glu	Ser	Ser	Asn	Asn	Val	Glu	Ser	Gly		
				230					235					240			
aat	tac	tca	gac	tct	gga	aat	tac	att	caa	cca	gga	aac	aat	gtt	gtc	945	
Asn	Tyr	Ser	Asp	Ser	Gly	Asn	Tyr	Ile	Gln	Pro	Gly	Asn	Asn	Val	Val		
			245					250					255				
gac	tct	ggt	ggg	tac	ttt	gaa	caa	cca	att	gaa	act	ttt	gag	gaa	gat	993	
Asp	Ser	Gly	Gly	Tyr	Phe	Glu	Gln	Pro	Ile	Glu	Thr	Phe	Glu	Glu	Asp		
		260				265						270					
cgg	aag	cct	att	ata	cgg	gag	ggt	agc	att	cag	cct	tgt	tct	ctg	ttt	1041	
Arg	Lys	Pro	Ile	Ile	Arg	Glu	Gly	Ser	Ile	Gln	Pro	Cys	Ser	Leu	Phe		
	275					280					285						
cca	gag	gaa	caa	att	ggc	tgt	ggt	gtg	caa	gac	gaa	aat	gtg	gtg	aat	1089	
Pro	Glu	Glu	Gln	Ile	Gly	Cys	Gly	Val	Gln	Asp	Glu	Asn	Val	Val	Asn		
290					295				300					305			
ctg	gaa	tct	tcc	aac	aat	aat	gtg	ttt	gta	gct	gat	aca	tgc	tac	agt	1137	
Leu	Glu	Ser	Ser	Asn	Asn	Asn	Val	Phe	Val	Ala	Asp	Thr	Cys	Tyr	Ser		
				310				315						320			
gac	att	cct	att	gat	cat	aac	tat	tta	ccc	gat	gag	cca	ttc	atg	gat	1185	
Asp	Ile	Pro	Ile	Asp	His	Asn	Tyr	Leu	Pro	Asp	Glu	Pro	Phe	Met	Asp		
			325					330					335				
cct	aat	aac	aat	ctt	cca	ctc	aac	gat	ggt	ctg	tac	ctg	gaa	acg	aat	1233	
Pro	Asn	Asn	Asn	Leu	Pro	Leu	Asn	Asp	Gly	Leu	Tyr	Leu	Glu	Thr	Asn		
			340				345					350					
gat	ctc	agc	tgt	gct	caa	caa	gat	gat	ttt	aac	ttc	gaa	gat	tat	ctc	1281	
Asp	Leu	Ser	Cys	Ala	Gln	Gln	Asp	Asp	Phe	Asn	Phe	Glu	Asp	Tyr	Leu		
	355				360						365						
agc	ttc	ttt	gat	gat	gag	ggt	ttg	act	ttt	gac	gat	tct	cta	tta	atg	1329	
Ser	Phe	Phe	Asp	Asp	Glu	Gly	Leu	Thr	Phe	Asp	Asp	Ser	Leu	Leu	Met		

370	375	380	385	
gga cct gaa gat ttt ctt ccc aac caa gaa gcc ctt gac cag aaa cct				1377
Gly Pro Glu Asp Phe Leu Pro Asn Gln Glu Ala Leu Asp Gln Lys Pro				
	390	395	400	
gcc cct aaa gaa ttg gag aag gag gtc gca gga ggc aaa gag gca gtg				1425
Ala Pro Lys Glu Leu Glu Lys Glu Val Ala Gly Gly Lys Glu Ala Val				
	405	410	415	
gag gaa aag gaa agt ggc gaa gga tct tct tca aaa caa gat aca gat				1473
Glu Glu Lys Glu Ser Gly Glu Gly Ser Ser Ser Lys Gln Asp Thr Asp				
	420	425	430	
ttc aag gac ttt gat tca gct ccg aag tac cca ttt ctc aaa aag acg				1521
Phe Lys Asp Phe Asp Ser Ala Pro Lys Tyr Pro Phe Leu Lys Lys Thr				
	435	440	445	
agc cac atg ctt gga gcc att cct act cca tct tca ttt gct tca cag				1569
Ser His Met Leu Gly Ala Ile Pro Thr Pro Ser Ser Phe Ala Ser Gln				
	450	455	460	465
ttc caa aca aag gac gca atg cgt cta cac gca gca caa tct tct ggt				1617
Phe Gln Thr Lys Asp Ala Met Arg Leu His Ala Ala Gln Ser Ser Gly				
	470	475	480	
tca gtt cac gtg act gca ggt atg atg aga ata tca aac atg act cta				1665
Ser Val His Val Thr Ala Gly Met Met Arg Ile Ser Asn Met Thr Leu				
	485	490	495	
gca gcg gac agc ggt atg ggc tgg tca tat gac aag aac ggt aac ctc				1713
Ala Ala Asp Ser Gly Met Gly Trp Ser Tyr Asp Lys Asn Gly Asn Leu				
	500	505	510	
aac gta gtc ctt tca ttc ggg gta gtc caa cag gat gat gcg atg act				1761
Asn Val Val Leu Ser Phe Gly Val Val Gln Gln Asp Asp Ala Met Thr				
	515	520	525	
gcc tcg gga agc aag aca gga att acg gcg aca aga gct atg tta gtc				1809
Ala Ser Gly Ser Lys Thr Gly Ile Thr Ala Thr Arg Ala Met Leu Val				
	530	535	540	545
ttc atg tgt tta tgg gtt ctc cta ctc tct gtt agc ttc aaa ata gta				1857
Phe Met Cys Leu Trp Val Leu Leu Leu Ser Val Ser Phe Lys Ile Val				
	550	555	560	
acc atg gtg tct gct cgg taa taggatcaaa gttgaatcgt ctcaaagact				1908
Thr Met Val Ser Ala Arg *				
	565			
ttttttggtg tttgtacctc tccaatcata tagcctttaa ctttggcagt gctttgctgc				1968
tcaatattta aatttttaaaa aaaaaaaaaa aaa				2001

&lt;210&gt; 162

&lt;211&gt; 567

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana



&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (12)...(156)

&lt;223&gt; Conserved domain

&lt;400&gt; 162

```

Met Gly Arg Gly Ser Val Thr Ser Leu Ala Pro Gly Phe Arg Phe His
 1           5           10           15
Pro Thr Asp Glu Glu Leu Val Arg Tyr Tyr Leu Lys Arg Lys Val Cys
          20           25           30
Asn Lys Pro Phe Lys Phe Asp Ala Ile Ser Val Thr Asp Ile Tyr Lys
          35           40           45
Ser Glu Pro Trp Asp Leu Pro Asp Lys Ser Lys Leu Lys Ser Arg Asp
          50           55           60
Leu Glu Trp Tyr Phe Phe Ser Met Leu Asp Lys Lys Tyr Ser Asn Gly
          65           70           75           80
Ser Lys Thr Asn Arg Ala Thr Glu Lys Gly Tyr Trp Lys Thr Thr Gly
          85           90           95
Lys Asp Arg Glu Ile Arg Asn Gly Ser Arg Val Val Gly Met Lys Lys
          100          105          110
Thr Leu Val Tyr His Lys Gly Arg Ala Pro Arg Gly Glu Arg Thr Asn
          115          120          125
Trp Val Met His Glu Tyr Arg Leu Ser Asp Glu Asp Leu Lys Lys Ala
          130          135          140
Gly Val Pro Gln Glu Ala Tyr Val Leu Cys Arg Ile Phe Gln Lys Ser
          145          150          155          160
Gly Thr Gly Pro Lys Asn Gly Glu Gln Tyr Gly Ala Pro Tyr Leu Glu
          165          170          175
Glu Glu Trp Glu Glu Asp Gly Met Thr Tyr Val Pro Ala Gln Asp Ala
          180          185          190
Phe Ser Glu Gly Leu Ala Leu Asn Asp Asp Val Tyr Val Asp Ile Asp
          195          200          205
Asp Ile Asp Glu Lys Pro Glu Asn Leu Val Val Tyr Asp Ala Val Pro
          210          215          220
Ile Leu Pro Asn Tyr Cys His Gly Glu Ser Ser Asn Asn Val Glu Ser
          225          230          235          240
Gly Asn Tyr Ser Asp Ser Gly Asn Tyr Ile Gln Pro Gly Asn Asn Val
          245          250          255
Val Asp Ser Gly Gly Tyr Phe Glu Gln Pro Ile Glu Thr Phe Glu Glu
          260          265          270
Asp Arg Lys Pro Ile Ile Arg Glu Gly Ser Ile Gln Pro Cys Ser Leu
          275          280          285
Phe Pro Glu Glu Gln Ile Gly Cys Gly Val Gln Asp Glu Asn Val Val
          290          295          300
Asn Leu Glu Ser Ser Asn Asn Asn Val Phe Val Ala Asp Thr Cys Tyr
          305          310          315          320
Ser Asp Ile Pro Ile Asp His Asn Tyr Leu Pro Asp Glu Pro Phe Met
          325          330          335
Asp Pro Asn Asn Asn Leu Pro Leu Asn Asp Gly Leu Tyr Leu Glu Thr
          340          345          350
Asn Asp Leu Ser Cys Ala Gln Gln Asp Asp Phe Asn Phe Glu Asp Tyr
          355          360          365
Leu Ser Phe Phe Asp Asp Glu Gly Leu Thr Phe Asp Asp Ser Leu Leu
          370          375          380
Met Gly Pro Glu Asp Phe Leu Pro Asn Gln Glu Ala Leu Asp Gln Lys
          385          390          395          400

```

Pro Ala Pro Lys Glu Leu Glu Lys Glu Val Ala Gly Gly Lys Glu Ala  
 405 410 415  
 Val Glu Glu Lys Glu Ser Gly Glu Gly Ser Ser Ser Lys Gln Asp Thr  
 420 425 430  
 Asp Phe Lys Asp Phe Asp Ser Ala Pro Lys Tyr Pro Phe Leu Lys Lys  
 435 440 445  
 Thr Ser His Met Leu Gly Ala Ile Pro Thr Pro Ser Ser Phe Ala Ser  
 450 455 460  
 Gln Phe Gln Thr Lys Asp Ala Met Arg Leu His Ala Ala Gln Ser Ser  
 465 470 475 480  
 Gly Ser Val His Val Thr Ala Gly Met Met Arg Ile Ser Asn Met Thr  
 485 490 495  
 Leu Ala Ala Asp Ser Gly Met Gly Trp Ser Tyr Asp Lys Asn Gly Asn  
 500 505 510  
 Leu Asn Val Val Leu Ser Phe Gly Val Val Gln Gln Asp Asp Ala Met  
 515 520 525  
 Thr Ala Ser Gly Ser Lys Thr Gly Ile Thr Ala Thr Arg Ala Met Leu  
 530 535 540  
 Val Phe Met Cys Leu Trp Val Leu Leu Leu Ser Val Ser Phe Lys Ile  
 545 550 555 560  
 Val Thr Met Val Ser Ala Arg  
 565

<210> 163  
 <211> 890  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (108)...(806)

<400> 163  
 cattcaaaaa catcatatat atacacaaac acactttgat acaacaaaaaaa aaaacagaac 60  
 acaacaaaaa acacattgta acattagttt aagcattaag cttcttt atg tcg aat 116  
 Met Ser Asn  
 1  
 aat aat aat tct ccg acc acc gtg aat caa gaa acg acg acg tct cgt 164  
 Asn Asn Asn Ser Pro Thr Thr Val Asn Gln Glu Thr Thr Thr Ser Arg  
 5 10 15  
 gaa gtc tca atc aca ttg cct act gat caa tct cct caa acc tca cca 212  
 Glu Val Ser Ile Thr Leu Pro Thr Asp Gln Ser Pro Gln Thr Ser Pro  
 20 25 30 35  
 gga tca tct tct tct cct tca ccg aga cct tcc ggt gga tca ccg gcg 260  
 Gly Ser Ser Ser Ser Pro Ser Pro Arg Pro Ser Gly Gly Ser Pro Ala  
 40 45 50  
 aga aga acg gcg act gga tta tcc ggc aag cac tct att ttc agg ggg 308  
 Arg Arg Thr Ala Thr Gly Leu Ser Gly Lys His Ser Ile Phe Arg Gly  
 55 60 65  
 att cga cta cgt aac gga aaa tgg gta tcg gag att aga gag cca cgt 356  
 Ile Arg Leu Arg Asn Gly Lys Trp Val Ser Glu Ile Arg Glu Pro Arg

70	75	80	
aaa acg aca aga att tgg ctc ggg act tat ccg gta ccg gag atg gct			404
Lys Thr Thr Arg Ile Trp Leu Gly Thr Tyr Pro Val Pro Glu Met Ala			
85	90	95	
gcc gcc gct tac gac gtg gct gcg tta gct tta aaa gga ccc gac gcc			452
Ala Ala Ala Tyr Asp Val Ala Ala Leu Ala Leu Lys Gly Pro Asp Ala			
100	105	110	115
gtt ttg aat ttt cct ggt tta gct ttg act tac gtg gct ccg gtt tca			500
Val Leu Asn Phe Pro Gly Leu Ala Leu Thr Tyr Val Ala Pro Val Ser			
120	125	130	
aac tct gct gcg gat ata aga gcg gct gct agt aga gca gcg gag atg			548
Asn Ser Ala Ala Asp Ile Arg Ala Ala Ala Ser Arg Ala Ala Glu Met			
135	140	145	
aag caa ccg gat cag ggt ggg gat gag aag gta ttg gaa ccg gtt caa			596
Lys Gln Pro Asp Gln Gly Gly Asp Glu Lys Val Leu Glu Pro Val Gln			
150	155	160	
ccc ggc aaa gag gaa gaa tta gaa gaa gtg tcg tgt aac tcg tgt tcg			644
Pro Gly Lys Glu Glu Glu Leu Glu Glu Val Ser Cys Asn Ser Cys Ser			
165	170	175	
ttg gag ttt atg gat gag gaa gcg atg ttg aat atg ccg act ttg ttg			692
Leu Glu Phe Met Asp Glu Glu Ala Met Leu Asn Met Pro Thr Leu Leu			
180	185	190	195
acg gag atg gct gaa ggg atg ttg atg agt cca ccg aga atg atg ata			740
Thr Glu Met Ala Glu Gly Met Leu Met Ser Pro Pro Arg Met Met Ile			
200	205	210	
cat ccg acg atg gaa gat gat tcg ccg gag aat cat gaa gga gat aat			788
His Pro Thr Met Glu Asp Asp Ser Pro Glu Asn His Glu Gly Asp Asn			
215	220	225	
ctt tgg agt tat aaa tga atccattgaa gctgctctct tttttattgt			836
Leu Trp Ser Tyr Lys *			
230			
tttccggtcg aatgagattt tccccctttt tttttttctt tttgggtcgc tggt			890
<210> 164			
<211> 232			
<212> PRT			
<213> Arabidopsis thaliana			
<220>			
<221> DOMAIN			
<222> (62)...(128)			
<223> Conserved domain			
<400> 164			
Met Ser Asn Asn Asn Asn Ser Pro Thr Thr Val Asn Gln Glu Thr Thr			
1	5	10	15

Thr Ser Arg Glu Val Ser Ile Thr Leu Pro Thr Asp Gln Ser Pro Gln  
 20 25 30  
 Thr Ser Pro Gly Ser Ser Ser Ser Pro Ser Pro Arg Pro Ser Gly Gly  
 35 40 45  
 Ser Pro Ala Arg Arg Thr Ala Thr Gly Leu Ser Gly Lys His Ser Ile  
 50 55 60  
 Phe Arg Gly Ile Arg Leu Arg Asn Gly Lys Trp Val Ser Glu Ile Arg  
 65 70 75 80  
 Glu Pro Arg Lys Thr Thr Arg Ile Trp Leu Gly Thr Tyr Pro Val Pro  
 85 90 95  
 Glu Met Ala Ala Ala Tyr Asp Val Ala Ala Leu Ala Leu Lys Gly  
 100 105 110  
 Pro Asp Ala Val Leu Asn Phe Pro Gly Leu Ala Leu Thr Tyr Val Ala  
 115 120 125  
 Pro Val Ser Asn Ser Ala Ala Asp Ile Arg Ala Ala Ser Arg Ala  
 130 135 140  
 Ala Glu Met Lys Gln Pro Asp Gln Gly Gly Asp Glu Lys Val Leu Glu  
 145 150 155 160  
 Pro Val Gln Pro Gly Lys Glu Glu Glu Leu Glu Glu Val Ser Cys Asn  
 165 170 175  
 Ser Cys Ser Leu Glu Phe Met Asp Glu Glu Ala Met Leu Asn Met Pro  
 180 185 190  
 Thr Leu Leu Thr Glu Met Ala Glu Gly Met Leu Met Ser Pro Pro Arg  
 195 200 205  
 Met Met Ile His Pro Thr Met Glu Asp Asp Ser Pro Glu Asn His Glu  
 210 215 220  
 Gly Asp Asn Leu Trp Ser Tyr Lys  
 225 230

&lt;210&gt; 165

&lt;211&gt; 1728

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (106)...(1575)

&lt;400&gt; 165

cttttattttt tttccctct tgcgaaaac atcacatata aataaaaaac caaaaaccct 60  
 aatttataaa atatataaa aatagttata gctcccccact tatat atg caa aag cca 117  
 Met Gln Lys Pro  
 1

aca tca agt atc tta aat gtc ata atg gac ggt gga gac agc gtc gga 165  
 Thr Ser Ser Ile Leu Asn Val Ile Met Asp Gly Gly Asp Ser Val Gly  
 5 10 15 20

gga gga gga gga gat gat cac cac cgt cac ctc cac cat cac cac cgc 213  
 Gly Gly Gly Gly Asp Asp His His Arg His Leu His His His His Arg  
 25 30 35

cct act ttc cct ttt caa cta ctc gga aaa cac gac ccc gac gac aac 261  
 Pro Thr Phe Pro Phe Gln Leu Leu Gly Lys His Asp Pro Asp Asp Asn  
 40 45 50

cat caa caa caa cct tcc cct tcc tcc tcc tcc tct cta ttt tct ctc	309
His Gln Gln Gln Pro Ser Pro Ser Ser Ser Ser Ser Ser Leu Phe Ser Leu	
55 60 65	
cat caa cac caa caa tta tct caa tcg caa cct caa tcg caa tcg caa	357
His Gln His Gln Gln Leu Ser Gln Ser Gln Pro Gln Ser Gln Ser Gln	
70 75 80	
aag tca caa ccg cag aca acg caa aaa gag tta tta caa acg caa gag	405
Lys Ser Gln Pro Gln Thr Thr Gln Lys Glu Leu Leu Gln Thr Gln Glu	
85 90 95 100	
gaa tct gcg gtg gtg gca gct aaa aag cca ccg ttg aaa cga gcg tcg	453
Glu Ser Ala Val Val Ala Ala Lys Lys Pro Pro Leu Lys Arg Ala Ser	
105 110 115	
acg aaa gac cga cac acg aaa gta gac gga aga ggg agg aga ata agg	501
Thr Lys Asp Arg His Thr Lys Val Asp Gly Arg Gly Arg Arg Ile Arg	
120 125 130	
atg ccg gcg tta tgt gca gct agg gtt ttt cag cta acg cga gag cta	549
Met Pro Ala Leu Cys Ala Ala Arg Val Phe Gln Leu Thr Arg Glu Leu	
135 140 145	
ggc cat aaa tcc gac ggt gag aca ata gag tgg ctt ctt caa caa gct	597
Gly His Lys Ser Asp Gly Glu Thr Ile Glu Trp Leu Leu Gln Gln Ala	
150 155 160	
gaa cca tct gta atc gcc gcc acc gga acc gga aca atc ccg gcg aat	645
Glu Pro Ser Val Ile Ala Ala Thr Gly Thr Gly Thr Ile Pro Ala Asn	
165 170 175 180	
ttc act tct tta aac atc tct ctc cgt tct tca ggc tct tcc atg tct	693
Phe Thr Ser Leu Asn Ile Ser Leu Arg Ser Ser Gly Ser Ser Met Ser	
185 190 195	
ctt cct tct cat ttc cgc tcc gcc gct tcc act ttt agc cct aat aac	741
Leu Pro Ser His Phe Arg Ser Ala Ala Ser Thr Phe Ser Pro Asn Asn	
200 205 210	
ata ttt tct ccg gcg atg ctt caa caa caa caa caa caa caa cgt ggt	789
Ile Phe Ser Pro Ala Met Leu Gln Gln Gln Gln Gln Gln Gln Arg Gly	
215 220 225	
ggc ggt gtt ggg ttt cat cat ccc cat cta cag gga cgt gca cct acg	837
Gly Gly Val Gly Phe His His Pro His Leu Gln Gly Arg Ala Pro Thr	
230 235 240	
tcg tct ttg ttt cct ggt att gat aac ttc aca cca acg acg tcg ttt	885
Ser Ser Leu Phe Pro Gly Ile Asp Asn Phe Thr Pro Thr Thr Ser Phe	
245 250 255 260	
ttg aac ttt cat aat cca aca aag caa gaa gga gat caa gat tct gaa	933
Leu Asn Phe His Asn Pro Thr Lys Gln Glu Gly Asp Gln Asp Ser Glu	
265 270 275	
gag tta aac tcg gag aag aaa aga aga atc caa acg acg tcg gat ttg	981

Glu Leu Asn Ser Glu Lys Lys Arg Arg Ile Gln Thr Thr Ser Asp Leu	
280 285 290	
cat caa caa caa caa caa cac caa cat gat caa atc gga gga tat aca	1029
His Gln Gln Gln Gln Gln His Gln His Asp Gln Ile Gly Gly Tyr Thr	
295 300 305	
ctt caa tct agc aac agt gga tct acg gcc acc gca gcc gcc gcg caa	1077
Leu Gln Ser Ser Asn Ser Gly Ser Thr Ala Thr Ala Ala Ala Ala Gln	
310 315 320	
caa ata ccg gga aat ttc tgg atg gtt gcg gcg gct gcg gct gca ggt	1125
Gln Ile Pro Gly Asn Phe Trp Met Val Ala Ala Ala Ala Ala Ala Gly	
325 330 335 340	
ggt ggt ggt ggt aat aac aac caa aca ggt ggt ctt atg aca gct tct	1173
Gly Gly Gly Gly Asn Asn Asn Gln Thr Gly Gly Leu Met Thr Ala Ser	
345 350 355	
att ggt act ggt ggc ggc ggt gga gag cct gtt tgg acg ttt cct tcc	1221
Ile Gly Thr Gly Gly Gly Gly Gly Glu Pro Val Trp Thr Phe Pro Ser	
360 365 370	
att aac acg gca gcg gca gcg tta tat aga agt ggc gtt tcg ggc gtt	1269
Ile Asn Thr Ala Ala Ala Ala Leu Tyr Arg Ser Gly Val Ser Gly Val	
375 380 385	
cca agc ggc gcg gtt tct agc ggt tta cat ttt atg aat ttc gca gcg	1317
Pro Ser Gly Ala Val Ser Ser Gly Leu His Phe Met Asn Phe Ala Ala	
390 395 400	
cca atg gca ttt ctt act gga caa caa cag cta gca aca act agt aat	1365
Pro Met Ala Phe Leu Thr Gly Gln Gln Gln Ala Thr Thr Ser Asn	
405 410 415 420	
cat gag att aat gaa gat agt aat aat aat gaa gga gga aga agt gac	1413
His Glu Ile Asn Glu Asp Ser Asn Asn Asn Glu Gly Gly Arg Ser Asp	
425 430 435	
ggt ggt ggt gat cat cat aat aca cag aga cat cat cat cat caa caa	1461
Gly Gly Gly Asp His His Asn Thr Gln Arg His His His His Gln Gln	
440 445 450	
caa cat cat cat aat att ctc tcc ggc ttg aac cag tac gga cgg caa	1509
Gln His His His Asn Ile Leu Ser Gly Leu Asn Gln Tyr Gly Arg Gln	
455 460 465	
gtt tcc ggc gac tct caa gct agt gga tca ctt gga ggt ggt gat gag	1557
Val Ser Gly Asp Ser Gln Ala Ser Gly Ser Leu Gly Gly Gly Asp Glu	
470 475 480	
gag gat cag caa gat tag acacacacaa aaacatttta atggtgggat	1605
Glu Asp Gln Gln Asp *	
485	
tttctgcca cggcggtagc ggtgacggcg ttcggcggt gtgtaaaact tttgttttca	1665
ccgttaattt tttttatttt ctgtagtaat ttttttgct aagtaaaaaa aaaaaaaaaa	1725

aaa

1728

&lt;210&gt; 166

&lt;211&gt; 489

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (116)...(179)

&lt;223&gt; Conserved domain

&lt;400&gt; 166

```

Met Gln Lys Pro Thr Ser Ser Ile Leu Asn Val Ile Met Asp Gly Gly
 1          5          10          15
Asp Ser Val Gly Gly Gly Gly Asp Asp His His Arg His Leu His
 20          25          30
His His His Arg Pro Thr Phe Pro Phe Gln Leu Leu Gly Lys His Asp
 35          40          45
Pro Asp Asp Asn His Gln Gln Gln Pro Ser Pro Ser Ser Ser Ser
 50          55          60
Leu Phe Ser Leu His Gln His Gln Gln Leu Ser Gln Ser Gln Pro Gln
 65          70          75          80
Ser Gln Ser Gln Lys Ser Gln Pro Gln Thr Thr Gln Lys Glu Leu Leu
 85          90          95
Gln Thr Gln Glu Glu Ser Ala Val Val Ala Ala Lys Lys Pro Pro Leu
100          105          110
Lys Arg Ala Ser Thr Lys Asp Arg His Thr Lys Val Asp Gly Arg Gly
115          120          125
Arg Arg Ile Arg Met Pro Ala Leu Cys Ala Ala Arg Val Phe Gln Leu
130          135          140
Thr Arg Glu Leu Gly His Lys Ser Asp Gly Glu Thr Ile Glu Trp Leu
145          150          155          160
Leu Gln Gln Ala Glu Pro Ser Val Ile Ala Ala Thr Gly Thr Gly Thr
165          170          175
Ile Pro Ala Asn Phe Thr Ser Leu Asn Ile Ser Leu Arg Ser Ser Gly
180          185          190
Ser Ser Met Ser Leu Pro Ser His Phe Arg Ser Ala Ala Ser Thr Phe
195          200          205
Ser Pro Asn Asn Ile Phe Ser Pro Ala Met Leu Gln Gln Gln Gln Gln
210          215          220
Gln Gln Arg Gly Gly Gly Val Gly Phe His His Pro His Leu Gln Gly
225          230          235          240
Arg Ala Pro Thr Ser Ser Leu Phe Pro Gly Ile Asp Asn Phe Thr Pro
245          250          255
Thr Thr Ser Phe Leu Asn Phe His Asn Pro Thr Lys Gln Glu Gly Asp
260          265          270
Gln Asp Ser Glu Glu Leu Asn Ser Glu Lys Lys Arg Arg Ile Gln Thr
275          280          285
Thr Ser Asp Leu His Gln Gln Gln Gln Gln His Gln His Asp Gln Ile
290          295          300
Gly Gly Tyr Thr Leu Gln Ser Ser Asn Ser Gly Ser Thr Ala Thr Ala
305          310          315          320
Ala Ala Ala Gln Gln Ile Pro Gly Asn Phe Trp Met Val Ala Ala Ala
325          330          335
Ala Ala Ala Gly Gly Gly Gly Gly Asn Asn Asn Gln Thr Gly Gly Leu
340          345          350

```

Met Thr Ala Ser Ile Gly Thr Gly Gly Gly Gly Gly Glu Pro Val Trp  
 355 360 365  
 Thr Phe Pro Ser Ile Asn Thr Ala Ala Ala Ala Leu Tyr Arg Ser Gly  
 370 375 380  
 Val Ser Gly Val Pro Ser Gly Ala Val Ser Ser Gly Leu His Phe Met  
 385 390 395 400  
 Asn Phe Ala Ala Pro Met Ala Phe Leu Thr Gly Gln Gln Gln Leu Ala  
 405 410 415  
 Thr Thr Ser Asn His Glu Ile Asn Glu Asp Ser Asn Asn Asn Glu Gly  
 420 425 430  
 Gly Arg Ser Asp Gly Gly Gly Asp His His Asn Thr Gln Arg His His  
 435 440 445  
 His His Gln Gln Gln His His His Asn Ile Leu Ser Gly Leu Asn Gln  
 450 455 460  
 Tyr Gly Arg Gln Val Ser Gly Asp Ser Gln Ala Ser Gly Ser Leu Gly  
 465 470 475 480  
 Gly Gly Asp Glu Glu Asp Gln Gln Asp  
 485

&lt;210&gt; 167

&lt;211&gt; 2535

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(2535)

&lt;400&gt; 167

atg gga tgt tct cat tcc aag tta gat gat gaa gaa gca gtt cag atc	48
Met Gly Cys Ser His Ser Lys Leu Asp Asp Glu Glu Ala Val Gln Ile	
1 5 10 15	
tgt aaa gac agg aaa cgt ttc atc aaa caa gct att gaa cat aga act	96
Cys Lys Asp Arg Lys Arg Phe Ile Lys Gln Ala Ile Glu His Arg Thr	
20 25 30	
aag ttt gct tct ggt cac att gct tat att cac tct ctt aga aaa gtc	144
Lys Phe Ala Ser Gly His Ile Ala Tyr Ile His Ser Leu Arg Lys Val	
35 40 45	
tct gat gct ctc cat gac ttt atc ctc caa gga gac aac aac aat gag	192
Ser Asp Ala Leu His Asp Phe Ile Leu Gln Gly Asp Asn Asn Asn Glu	
50 55 60	
ttt gtt cct act ctc tgt caa gac tca ttc gtc act ccg gtc aag aga	240
Phe Val Pro Thr Leu Cys Gln Asp Ser Phe Val Thr Pro Val Lys Arg	
65 70 75 80	
atg cca cag aga cgt cgt cga agc agc cgt agt aat agc ggt gag ttc	288
Met Pro Gln Arg Arg Arg Arg Ser Ser Arg Ser Asn Ser Gly Glu Phe	
85 90 95	
atc tcc att tcg cct tct tcg att cct ccg aaa atg att caa ggt agg	336
Ile Ser Ile Ser Pro Ser Ser Ile Pro Pro Lys Met Ile Gln Gly Arg	
100 105 110	



cca aga tcg aat gtg aaa gcg aat tat ttg atg gct aat aga agt aga	384
Pro Arg Ser Asn Val Lys Ala Asn Tyr Leu Met Ala Asn Arg Ser Arg	
115 120 125	
ccg gtt aga gtc gaa cag aga tct cct gaa acg ttt aga gtt gag aca	432
Pro Val Arg Val Glu Gln Arg Ser Pro Glu Thr Phe Arg Val Glu Thr	
130 135 140	
ttt tct cca cct agt agt agt aac caa tat ggt gaa tca gat ggt ttc	480
Phe Ser Pro Pro Ser Ser Ser Asn Gln Tyr Gly Glu Ser Asp Gly Phe	
145 150 155 160	
ttt gga atg aat atg aat atg aat act tct gcg tct act tca tct tcg	528
Phe Gly Met Asn Met Asn Met Asn Thr Ser Ala Ser Thr Ser Ser Ser	
165 170 175	
ttt tgg aat ccg ttg agc tct cct gag caa agg tta agt agt cat aac	576
Phe Trp Asn Pro Leu Ser Ser Pro Glu Gln Arg Leu Ser Ser His Asn	
180 185 190	
att cct cct cct tca ccg cag aat tcg cag tgg gat ttc ttc tgg aat	624
Ile Pro Pro Pro Ser Pro Gln Asn Ser Gln Trp Asp Phe Phe Trp Asn	
195 200 205	
ccg ttt tct tcg ttg gat tat tat gga tac aat agt tat gac cga gga	672
Pro Phe Ser Ser Leu Asp Tyr Tyr Gly Tyr Asn Ser Tyr Asp Arg Gly	
210 215 220	
agt gtt gat agt agg agt ggt ata gat gat gag att aga gga ttg aga	720
Ser Val Asp Ser Arg Ser Gly Ile Asp Asp Glu Ile Arg Gly Leu Arg	
225 230 235 240	
cgt gtt aga gag gaa gaa ggg att cct gat ttg gaa gaa gat gat gaa	768
Arg Val Arg Glu Glu Glu Gly Ile Pro Asp Leu Glu Glu Asp Asp Glu	
245 250 255	
cct cat aaa cct gaa cct gtc gtc aat gtg cgg ttt cag aat cat aat	816
Pro His Lys Pro Glu Pro Val Val Asn Val Arg Phe Gln Asn His Asn	
260 265 270	
ccg aat ccg aaa gca acg gaa gag aat aga ggt aaa gtt gat aag agt	864
Pro Asn Pro Lys Ala Thr Glu Glu Asn Arg Gly Lys Val Asp Lys Ser	
275 280 285	
tgt tgc aat gag gaa gtt aaa gtt gaa gat gtt gat gag gat gaa gac	912
Cys Cys Asn Glu Glu Val Lys Val Glu Asp Val Asp Glu Asp Glu Asp	
290 295 300	
gag gat gga gat gag gat gat gat gag ttt acg gat agc ggg tgt gag	960
Glu Asp Gly Asp Glu Asp Asp Asp Glu Phe Thr Asp Ser Gly Cys Glu	
305 310 315 320	
act gag aat gag gga gac gag aaa tgc gtt gct cct cca acg caa gaa	1008
Thr Glu Asn Glu Gly Asp Glu Lys Cys Val Ala Pro Pro Thr Gln Glu	
325 330 335	

cag cga aaa gct gag gtg tgt aga gga gga gaa gct aca ggg aat gtt	1056
Gln Arg Lys Ala Glu Val Cys Arg Gly Gly Glu Ala Thr Gly Asn Val	
340 345 350	
ggt gga gta ggt aag gtt caa gag atg aag aat gtt gtt ggt gtt aga	1104
Val Gly Val Gly Lys Val Gln Glu Met Lys Asn Val Val Gly Val Arg	
355 360 365	
gat gat gca aag acg acg ggg ttc acg gta tat gtg aac cgg aga cca	1152
Asp Asp Ala Lys Thr Thr Gly Phe Thr Val Tyr Val Asn Arg Arg Pro	
370 375 380	
act agc atg gct gaa gta atc aaa gat ctt gaa gat cag ttc aca aca	1200
Thr Ser Met Ala Glu Val Ile Lys Asp Leu Glu Asp Gln Phe Thr Thr	
385 390 395 400	
att tgt gat gcg gct aag gaa gtc tcg ggg cta ttg gaa gct ggt aga	1248
Ile Cys Asp Ala Ala Lys Glu Val Ser Gly Leu Leu Glu Ala Gly Arg	
405 410 415	
gct caa tac aca tca tcc ttt aat gac cat agt ggt gag ttt atg aac	1296
Ala Gln Tyr Thr Ser Ser Phe Asn Asp His Ser Gly Glu Phe Met Asn	
420 425 430	
aat gtt gca ata ttt ctc tcc agg aaa atg ctg aat cca gta gcg ttg	1344
Asn Val Ala Ile Phe Leu Ser Arg Lys Met Leu Asn Pro Val Ala Leu	
435 440 445	
ttc cgc tcg ggt tca tca aga tct tcc tcg tca aga ttc ttg atc act	1392
Phe Arg Ser Gly Ser Ser Arg Ser Ser Ser Ser Arg Phe Leu Ile Thr	
450 455 460	
tct tct ggt ggt tca aga gag agc ggg tct gaa agc aga agc gat gtt	1440
Ser Ser Gly Gly Ser Arg Glu Ser Gly Ser Glu Ser Arg Ser Asp Val	
465 470 475 480	
tca gat gaa tct tgc atg att tca ggt agc cac caa aca aca ttg gac	1488
Ser Asp Glu Ser Cys Met Ile Ser Gly Ser His Gln Thr Thr Leu Asp	
485 490 495	
aga ctt ttc gcg tgg gag aag aaa ctc tat gat gaa gtt agg gta agt	1536
Arg Leu Phe Ala Trp Glu Lys Lys Leu Tyr Asp Glu Val Arg Val Ser	
500 505 510	
aac aag aga tat aag caa cta tct gga gaa cgt gtt cga aga gca tat	1584
Asn Lys Arg Tyr Lys Gln Leu Ser Gly Glu Arg Val Arg Arg Ala Tyr	
515 520 525	
gaa aag aaa tgt atg cag ctg agg aat cag gac gtg aaa gga gat gat	1632
Glu Lys Lys Cys Met Gln Leu Arg Asn Gln Asp Val Lys Gly Asp Asp	
530 535 540	
ccg ctt gcg gtt gat aaa act agg gct act atc aga gat cta gat aca	1680
Pro Leu Ala Val Asp Lys Thr Arg Ala Thr Ile Arg Asp Leu Asp Thr	
545 550 555 560	
cag atc aag gtg tca ata cac tcc att gaa tcc atc tcc aaa aga atc	1728

Gln Ile Lys Val Ser Ile His Ser Ile Glu Ser Ile Ser Lys Arg Ile	
565 570 575	
gag act ctt cgt gac caa gaa ctg ttg cct cag ctt ctt gag ctc gtt	1776
Glu Thr Leu Arg Asp Gln Glu Leu Leu Pro Gln Leu Leu Glu Leu Val	
580 585 590	
gaa gga ctg aca aat ggt aat cat gat cat aac aga tta aca agg atg	1824
Glu Gly Leu Thr Asn Gly Asn His Asp His Asn Arg Leu Thr Arg Met	
595 600 605	
tgg aaa gta atg gcg gaa agt cac cag ata cag aaa cga aca cta gac	1872
Trp Lys Val Met Ala Glu Ser His Gln Ile Gln Lys Arg Thr Leu Asp	
610 615 620	
gaa gcc aaa ctg tta ctt gca ggc aca ccg gtt tca aag cgt cac aag	1920
Glu Ala Lys Leu Leu Leu Ala Gly Thr Pro Val Ser Lys Arg His Lys	
625 630 635 640	
aag aga caa cct cca ata atg cca gag gct att aac tct caa aga tta	1968
Lys Arg Gln Pro Ile Met Pro Glu Ala Ile Asn Ser Gln Arg Leu	
645 650 655	
gct caa tct gct ttg aat ctt gaa gct cag ctt cga aac tgg aga acc	2016
Ala Gln Ser Ala Leu Asn Leu Glu Ala Gln Leu Arg Asn Trp Arg Thr	
660 665 670	
tgt ttt gaa ttc tgg atc acc tct cag aga tcc tac atg aaa gca tta	2064
Cys Phe Glu Phe Trp Ile Thr Ser Gln Arg Ser Tyr Met Lys Ala Leu	
675 680 685	
tct ggt tgg ctt cta aga tgt ttc aga tgt gat cct gac ccc gaa aaa	2112
Ser Gly Trp Leu Leu Arg Cys Phe Arg Cys Asp Pro Asp Pro Glu Lys	
690 695 700	
gtt aga tta tcc tca tgt ctt cac ccg ata tac cga gtc tgc att caa	2160
Val Arg Leu Ser Ser Cys Leu His Pro Ile Tyr Arg Val Cys Ile Gln	
705 710 715 720	
tgg tca agg ttg ctt aac agc ttg aat gaa aaa cca gtt ttg gac aaa	2208
Trp Ser Arg Leu Leu Asn Ser Leu Asn Glu Lys Pro Val Leu Asp Lys	
725 730 735	
ctt gag ttc ttt gcc tca ggg atg ggt tca ata tat gca agg caa gtt	2256
Leu Glu Phe Phe Ala Ser Gly Met Gly Ser Ile Tyr Ala Arg Gln Val	
740 745 750	
agg gaa gat cca aac tgg agt gga agc ggg tca agg aga tac tcc ggg	2304
Arg Glu Asp Pro Asn Trp Ser Gly Ser Gly Ser Arg Arg Tyr Ser Gly	
755 760 765	
tca gaa agc atg gag ctt gta gtg gct gat aag gga gaa gaa gat gtt	2352
Ser Glu Ser Met Glu Leu Val Val Ala Asp Lys Gly Glu Glu Asp Val	
770 775 780	
gtg atg act gct gag aaa ctt gca gag gtt gct gtg aag gtt ctc tgc	2400
Val Met Thr Ala Glu Lys Leu Ala Glu Val Ala Val Lys Val Leu Cys	

785	790	795	800	
cat gga atg tct gtt gca gtt agt tca cta gct gag ttt tcc atc aac				2448
His Gly Met Ser Val Ala Val Ser Ser Leu Ala Glu Phe Ser Ile Asn				
805		810	815	
tcg gcc gat gaa cac tcg aag ctt gtt agc caa cca gca gac gat gcg				2496
Ser Ala Asp Glu His Ser Lys Leu Val Ser Gln Pro Ala Asp Asp Ala				
820	825		830	
tcg gag cag cag ccc ccg gat gta aat ttc aat tcc tag				2535
Ser Glu Gln Gln Pro Pro Asp Val Asn Phe Asn Ser *				
835	840			

&lt;210&gt; 168

&lt;211&gt; 844

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 168

Met Gly Cys Ser His Ser Lys Leu Asp Asp Glu Glu Ala Val Gln Ile				
1	5	10	15	
Cys Lys Asp Arg Lys Arg Phe Ile Lys Gln Ala Ile Glu His Arg Thr				
20	25	30		
Lys Phe Ala Ser Gly His Ile Ala Tyr Ile His Ser Leu Arg Lys Val				
35	40	45		
Ser Asp Ala Leu His Asp Phe Ile Leu Gln Gly Asp Asn Asn Asn Glu				
50	55	60		
Phe Val Pro Thr Leu Cys Gln Asp Ser Phe Val Thr Pro Val Lys Arg				
65	70	75	80	
Met Pro Gln Arg Arg Arg Arg Ser Ser Arg Ser Asn Ser Gly Glu Phe				
85	90	95		
Ile Ser Ile Ser Pro Ser Ser Ile Pro Pro Lys Met Ile Gln Gly Arg				
100	105	110		
Pro Arg Ser Asn Val Lys Ala Asn Tyr Leu Met Ala Asn Arg Ser Arg				
115	120	125		
Pro Val Arg Val Glu Gln Arg Ser Pro Glu Thr Phe Arg Val Glu Thr				
130	135	140		
Phe Ser Pro Pro Ser Ser Ser Asn Gln Tyr Gly Glu Ser Asp Gly Phe				
145	150	155	160	
Phe Gly Met Asn Met Asn Met Asn Thr Ser Ala Ser Thr Ser Ser				
165	170	175		
Phe Trp Asn Pro Leu Ser Ser Pro Glu Gln Arg Leu Ser Ser His Asn				
180	185	190		
Ile Pro Pro Pro Ser Pro Gln Asn Ser Gln Trp Asp Phe Phe Trp Asn				
195	200	205		
Pro Phe Ser Ser Leu Asp Tyr Tyr Gly Tyr Asn Ser Tyr Asp Arg Gly				
210	215	220		
Ser Val Asp Ser Arg Ser Gly Ile Asp Asp Glu Ile Arg Gly Leu Arg				
225	230	235	240	
Arg Val Arg Glu Glu Glu Gly Ile Pro Asp Leu Glu Glu Asp Asp Glu				
245	250	255		
Pro His Lys Pro Glu Pro Val Val Asn Val Arg Phe Gln Asn His Asn				
260	265	270		
Pro Asn Pro Lys Ala Thr Glu Glu Asn Arg Gly Lys Val Asp Lys Ser				
275	280	285		

Cys Cys Asn Glu Glu Val Lys Val Glu Asp Val Asp Glu Asp Glu Asp  
 290 295 300  
 Glu Asp Gly Asp Glu Asp Asp Asp Glu Phe Thr Asp Ser Gly Cys Glu  
 305 310 315 320  
 Thr Glu Asn Glu Gly Asp Glu Lys Cys Val Ala Pro Pro Thr Gln Glu  
 325 330 335  
 Gln Arg Lys Ala Glu Val Cys Arg Gly Gly Glu Ala Thr Gly Asn Val  
 340 345 350  
 Val Gly Val Gly Lys Val Gln Glu Met Lys Asn Val Val Gly Val Arg  
 355 360 365  
 Asp Asp Ala Lys Thr Thr Gly Phe Thr Val Tyr Val Asn Arg Arg Pro  
 370 375 380  
 Thr Ser Met Ala Glu Val Ile Lys Asp Leu Glu Asp Gln Phe Thr Thr  
 385 390 395 400  
 Ile Cys Asp Ala Ala Lys Glu Val Ser Gly Leu Leu Glu Ala Gly Arg  
 405 410 415  
 Ala Gln Tyr Thr Ser Ser Phe Asn Asp His Ser Gly Glu Phe Met Asn  
 420 425 430  
 Asn Val Ala Ile Phe Leu Ser Arg Lys Met Leu Asn Pro Val Ala Leu  
 435 440 445  
 Phe Arg Ser Gly Ser Ser Arg Ser Ser Ser Arg Phe Leu Ile Thr  
 450 455 460  
 Ser Ser Gly Gly Ser Arg Glu Ser Gly Ser Glu Ser Arg Ser Asp Val  
 465 470 475 480  
 Ser Asp Glu Ser Cys Met Ile Ser Gly Ser His Gln Thr Thr Leu Asp  
 485 490 495  
 Arg Leu Phe Ala Trp Glu Lys Lys Leu Tyr Asp Glu Val Arg Val Ser  
 500 505 510  
 Asn Lys Arg Tyr Lys Gln Leu Ser Gly Glu Arg Val Arg Arg Ala Tyr  
 515 520 525  
 Glu Lys Lys Cys Met Gln Leu Arg Asn Gln Asp Val Lys Gly Asp Asp  
 530 535 540  
 Pro Leu Ala Val Asp Lys Thr Arg Ala Thr Ile Arg Asp Leu Asp Thr  
 545 550 555 560  
 Gln Ile Lys Val Ser Ile His Ser Ile Glu Ser Ile Ser Lys Arg Ile  
 565 570 575  
 Glu Thr Leu Arg Asp Gln Glu Leu Leu Pro Gln Leu Leu Glu Leu Val  
 580 585 590  
 Glu Gly Leu Thr Asn Gly Asn His Asp His Asn Arg Leu Thr Arg Met  
 595 600 605  
 Trp Lys Val Met Ala Glu Ser His Gln Ile Gln Lys Arg Thr Leu Asp  
 610 615 620  
 Glu Ala Lys Leu Leu Leu Ala Gly Thr Pro Val Ser Lys Arg His Lys  
 625 630 635 640  
 Lys Arg Gln Pro Pro Ile Met Pro Glu Ala Ile Asn Ser Gln Arg Leu  
 645 650 655  
 Ala Gln Ser Ala Leu Asn Leu Glu Ala Gln Leu Arg Asn Trp Arg Thr  
 660 665 670  
 Cys Phe Glu Phe Trp Ile Thr Ser Gln Arg Ser Tyr Met Lys Ala Leu  
 675 680 685  
 Ser Gly Trp Leu Leu Arg Cys Phe Arg Cys Asp Pro Asp Pro Glu Lys  
 690 695 700  
 Val Arg Leu Ser Ser Cys Leu His Pro Ile Tyr Arg Val Cys Ile Gln  
 705 710 715 720  
 Trp Ser Arg Leu Leu Asn Ser Leu Asn Glu Lys Pro Val Leu Asp Lys  
 725 730 735  
 Leu Glu Phe Phe Ala Ser Gly Met Gly Ser Ile Tyr Ala Arg Gln Val

```

      740      745      750
Arg Glu Asp Pro Asn Trp. Ser Gly Ser Gly Ser Arg Arg Tyr Ser Gly
      755      760      765
Ser Glu Ser Met Glu Leu Val Val Ala Asp Lys Gly Glu Glu Asp Val
      770      775      780
Val Met Thr Ala Glu Lys Leu Ala Glu Val Ala Val Lys Val Leu Cys
      785      790      795      800
His Gly Met Ser Val Ala Val Ser Ser Leu Ala Glu Phe Ser Ile Asn
      805      810      815
Ser Ala Asp Glu His Ser Lys Leu Val Ser Gln Pro Ala Asp Asp Ala
      820      825      830
Ser Glu Gln Gln Pro Pro Asp Val Asn Phe Asn Ser
      835      840

```

&lt;210&gt; 169

&lt;211&gt; 2054

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (390)...(1793)

&lt;400&gt; 169

```

catctgttct gagattgtct tcttcttctt gaaatctttt tcatttgtgc aatgattttg 60
aagactaccc atctattcag ttctttctct tatgatttaa gtttcattta gttttttttt 120
ctcatccgcc attttcggtg actttttgaa gcatttctcg aattgacttt gttcttcaca 180
ttgattcatt agaatgttct ttcataaagt ttggatcttt tcttcagggc ttgattaatt 240
tcatatctat gatcttctct gtattgtttt tgatccaatc acttctcaaa atttgatcct 300
tgtcattgga ttttagattta ggggttttgt atccttgggg atttgaagat caaaaacaga 360
gtctttgagt gatacttctg ggaacaaa atg gct gca act gca ata gag cca 413
                                Met Ala Ala Thr Ala Ile Glu Pro
                                1                                5

```

```

tct tca tct ata agt ttc aca tct tct cac tta tca aac cct tct cct 461
Ser Ser Ser Ile Ser Phe Thr Ser Ser His Leu Ser Asn Pro Ser Pro
    10                15                20

```

```

gtt gtt act act tat cac tca gct gct aat ctt gaa gag ctc agc tct 509
Val Val Thr Thr Tyr His Ser Ala Ala Asn Leu Glu Glu Leu Ser Ser
    25                30                35                40

```

```

aac ttg gag cag ctt ctc act aat cca gat tgc gat tac act gac gca 557
Asn Leu Glu Gln Leu Leu Thr Asn Pro Asp Cys Asp Tyr Thr Asp Ala
    45                50                55

```

```

gag atc atc att gaa gaa gaa gct aac cct cgg aag ctt cgt aac tat 605
Glu Ile Ile Ile Glu Glu Glu Ala Asn Pro Arg Lys Leu Arg Asn Tyr
    60                65                70

```

```

gtt gag aag tca cta gta gag aat gtt ctt cct atc ctc tta gtt gcg 653
Val Glu Lys Ser Leu Val Glu Asn Val Leu Pro Ile Leu Leu Val Ala
    75                80                85

```

```

ttt cat tgt gat ttg aca cag ctt ctt gat caa tgc att gag aga gtg 701
Phe His Cys Asp Leu Thr Gln Leu Leu Asp Gln Cys Ile Glu Arg Val

```

90	95	100	
gcg aga tca gac tta gac aga ttc tgt atc gaa aag gag ctt cct tta			749
Ala Arg Ser Asp Leu Asp Arg Phe Cys Ile Glu Lys Glu Leu Pro Leu			
105	110	115	120
gaa gta ttg gaa aaa atc aaa cag ctt cga gtt aag tcg gtg aac ata			797
Glu Val Leu Glu Lys Ile Lys Gln Leu Arg Val Lys Ser Val Asn Ile			
	125	130	135
ccc gag gtg gag gat aaa tcg ata gag aga aca ggg aaa gta ctc aag			845
Pro Glu Val Glu Asp Lys Ser Ile Glu Arg Thr Gly Lys Val Leu Lys			
	140	145	150
gca ttg gat tca gat gat gta gaa ctc gtg aag ctt ctt ttg act gag			893
Ala Leu Asp Ser Asp Asp Val Glu Leu Val Lys Leu Leu Leu Thr Glu			
	155	160	165
tca gat ata act cta gac caa gcc aat ggt cta cat tat gca gtg gca			941
Ser Asp Ile Thr Leu Asp Gln Ala Asn Gly Leu His Tyr Ala Val Ala			
	170	175	180
tac agt gat ccg aaa gtt gtg aca cag gtt ctt gat cta gat atg gct			989
Tyr Ser Asp Pro Lys Val Val Thr Gln Val Leu Asp Leu Asp Met Ala			
	185	190	200
gat gtt aat ttc aga aat tcc agg ggg tat acg gtt ctt cat att gct			1037
Asp Val Asn Phe Arg Asn Ser Arg Gly Tyr Thr Val Leu His Ile Ala			
	205	210	215
gct atg cgt aga gag cca aca att atc ata cca ctt att caa aaa gga			1085
Ala Met Arg Arg Glu Pro Thr Ile Ile Ile Pro Leu Ile Gln Lys Gly			
	220	225	230
gct aat gct tca gat ttc acg ttt gat gga cgc agt gcg gta aat ata			1133
Ala Asn Ala Ser Asp Phe Thr Phe Asp Gly Arg Ser Ala Val Asn Ile			
	235	240	245
tgt agg aga ctc act agg ccg aaa gat tat cat acc aaa acc tca agg			1181
Cys Arg Arg Leu Thr Arg Pro Lys Asp Tyr His Thr Lys Thr Ser Arg			
	250	255	260
aaa gaa cct agt aaa tac cgc tta tgc atc gat atc ttg gaa agg gaa			1229
Lys Glu Pro Ser Lys Tyr Arg Leu Cys Ile Asp Ile Leu Glu Arg Glu			
	265	270	275
att aga agg aat cca ttg gtt agt ggg gat aca ccc act tgt tcc cat			1277
Ile Arg Arg Asn Pro Leu Val Ser Gly Asp Thr Pro Thr Cys Ser His			
	285	290	295
tcg atg ccc gag gat ctc caa atg agg ttg tta tac tta gaa aag cga			1325
Ser Met Pro Glu Asp Leu Gln Met Arg Leu Leu Tyr Leu Glu Lys Arg			
	300	305	310
gtg gga ctt gct cag ttg ttc ttc cca gca gaa gcc aat gtg gct atg			1373
Val Gly Leu Ala Gln Leu Phe Phe Pro Ala Glu Ala Asn Val Ala Met			
	315	320	325

```

gac gtt gct aat gtt gaa ggg aca agc gag tgc aca ggt ctt cta act 1421
Asp Val Ala Asn Val Glu Gly Thr Ser Glu Cys Thr Gly Leu Leu Thr
330 335 340

cca cct cca tca aat gat aca act gaa aac ttg ggt aaa gtc gat tta 1469
Pro Pro Pro Ser Asn Asp Thr Thr Glu Asn Leu Gly Lys Val Asp Leu
345 350 355 360

aat gaa acg cct tat gtg caa acg aaa aga atg ctt aca cgt atg aaa 1517
Asn Glu Thr Pro Tyr Val Gln Thr Lys Arg Met Leu Thr Arg Met Lys
365 370 375

gcc ctc atg aaa aca gtt gag aca ggt cgg aga tac ttc cca tct tgt 1565
Ala Leu Met Lys Thr Val Glu Thr Gly Arg Arg Tyr Phe Pro Ser Cys
380 385 390

tat gag gtt ctg gat aag tac atg gat cag tat atg gac gaa gaa atc 1613
Tyr Glu Val Leu Asp Lys Tyr Met Asp Gln Tyr Met Asp Glu Glu Ile
395 400 405

cct gat atg tcg tat ccc gag aaa ggc act gtg aaa gag aga aga cag 1661
Pro Asp Met Ser Tyr Pro Glu Lys Gly Thr Val Lys Glu Arg Arg Gln
410 415 420

aag agg atg aga tat aac gag ctg aag aac gac gtt aaa aaa gca tat 1709
Lys Arg Met Arg Tyr Asn Glu Leu Lys Asn Asp Val Lys Lys Ala Tyr
425 430 435 440

agc aaa gac aaa gtc gcg cgg tct tgt ctt tct tct tca tca cca gct 1757
Ser Lys Asp Lys Val Ala Arg Ser Cys Leu Ser Ser Ser Ser Pro Ala
445 450 455

tct tct ctt aga gaa gcc tta gag aat cca aca tga tcttgcccc 1803
Ser Ser Leu Arg Glu Ala Leu Glu Asn Pro Thr *
460 465

gataagtaat gctactagag atagttcctt ggaagaaatt agctatttta taggttttgg 1863
tttagaaaac agtggagatc ccatttttgt attcaggttt attacatggg ttaggtttga 1923
ttttgtatta aaggagatgg atttttgaaa tgtataaagc aaaacctttc ctttttgctt 1983
tgcttctttc tttaaaaaaa aaattgtaat atatggaagt gaaatcgaag caacgattga 2043
gcaaaaaaaa a 2054

```

&lt;210&gt; 170

&lt;211&gt; 467

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (179)...(254)

&lt;223&gt; Conserved domain

&lt;400&gt; 170

```

Met Ala Ala Thr Ala Ile Glu Pro Ser Ser Ser Ile Ser Phe Thr Ser
1 5 10 15
Ser His Leu Ser Asn Pro Ser Pro Val Val Thr Thr Tyr His Ser Ala

```



234

<210> 171  
 <211> 1538  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (138)...(1388)

<400> 171  
 cagctcaaac tctctaggac tacactaaat ctaacttttt gcagagagca aaagattcaa 60  
 taattgagat tgatctcaaa accaaagctc tcgtgctctt gtcgttgatg ttggttggt 120  
 agactttgta tacaatg atg aaa agt ttg gcg aat gct gtt gga gcg aag 170  
 Met Lys Ser Leu Ala Asn Ala Val Gly Ala Lys  
 1 5 10

acg gcg agg gct tgc gac agc tgc gtg aag aga cgt gca cgg tgg tac 218  
 Thr Ala Arg Ala Cys Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr  
 15 20 25

tgc gcg gcc gac gat gct ttt ctt tgc cag tct tgc gac agt ttg gtc 266  
 Cys Ala Ala Asp Asp Ala Phe Leu Cys Gln Ser Cys Asp Ser Leu Val  
 30 35 40

cat tca gca aac cct ctt gct cgc cgc cac gag aga gtc cgt ttg aag 314  
 His Ser Ala Asn Pro Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys  
 45 50 55

acg gct agc ccg gcg gtc gta aag cat agc aac cac tca tca gct tct 362  
 Thr Ala Ser Pro Ala Val Val Lys His Ser Asn His Ser Ser Ala Ser  
 60 65 70 75

cct cca cat gag gtc gcc acg tgg cat cac ggg ttt act cgt aaa gct 410  
 Pro Pro His Glu Val Ala Thr Trp His His Gly Phe Thr Arg Lys Ala  
 80 85 90

cga acg cca cgt ggc tct ggt aag aaa aac aat tcg tcg ata ttt cat 458  
 Arg Thr Pro Arg Gly Ser Gly Lys Lys Asn Asn Ser Ser Ile Phe His  
 95 100 105

gac ttg gtt cct gat att agt att gag gat cag aca gac aac tat gag 506  
 Asp Leu Val Pro Asp Ile Ser Ile Glu Asp Gln Thr Asp Asn Tyr Glu  
 110 115 120

ctt gaa gag cag ctg atc tgt caa gtg ccg gtt cta gat ccg ttg gtg 554  
 Leu Glu Glu Gln Leu Ile Cys Gln Val Pro Val Leu Asp Pro Leu Val  
 125 130 135

tct gag cag ttc ttg aac gat gtc gtt gag ccc aag atc gag ttt cct 602  
 Ser Glu Gln Phe Leu Asn Asp Val Val Glu Pro Lys Ile Glu Phe Pro  
 140 145 150 155

atg atc aga agt ggt ttg atg atc gag gag gag gaa gac aac gct gaa 650  
 Met Ile Arg Ser Gly Leu Met Ile Glu Glu Glu Glu Asp Asn Ala Glu  
 160 165 170

```

agt tgt ctt aat gga ttt ttc ccg acc gac atg gag ctt gag gag ttt 698
Ser Cys Leu Asn Gly Phe Phe Pro Thr Asp Met Glu Leu Glu Glu Phe
175 180 185

gct gct gac gtg gag act ctg ctc ggt cgc ggg tta gac acg gag tcg 746
Ala Ala Asp Val Glu Thr Leu Leu Gly Arg Gly Leu Asp Thr Glu Ser
190 195 200

tat gcc atg gag gag cta ggg tta tct aat tca gag atg ttc aaa atc 794
Tyr Ala Met Glu Glu Leu Gly Leu Ser Asn Ser Glu Met Phe Lys Ile
205 210 215

gaa aaa gat gag att gaa gaa gaa gta gaa gag ata aaa gcc atg agc 842
Glu Lys Asp Glu Ile Glu Glu Glu Val Glu Glu Ile Lys Ala Met Ser
220 225 230 235

atg gat ata ttt gat gat gat cga aaa gac gtg gat gga aca gta ccg 890
Met Asp Ile Phe Asp Asp Asp Arg Lys Asp Val Asp Gly Thr Val Pro
240 245 250

ttt gag cta agc ttt gat tac gag tcg tca cac aag acg tcc gaa gaa 938
Phe Glu Leu Ser Phe Asp Tyr Glu Ser Ser His Lys Thr Ser Glu Glu
255 260 265

gag gta atg aag aac gtt gaa agt agt ggt gaa tgt gtt gtt aag gtg 986
Glu Val Met Lys Asn Val Glu Ser Ser Gly Glu Cys Val Val Lys Val
270 275 280

aaa gag gaa gaa cat aag aat gtt ctg atg cta aga tta aac tat gac 1034
Lys Glu Glu Glu His Lys Asn Val Leu Met Leu Arg Leu Asn Tyr Asp
285 290 295

tcg gtg ata tcc act tgg gga ggt caa ggt cca ccg tgg agt tca gga 1082
Ser Val Ile Ser Thr Trp Gly Gly Gln Gly Pro Pro Trp Ser Ser Gly
300 305 310 315

gag cca ccg gaa cga gac atg gac atc agc ggt tgg cca gcc ttt tcc 1130
Glu Pro Pro Glu Arg Asp Met Asp Ile Ser Gly Trp Pro Ala Phe Ser
320 325 330

atg gtg gag aat gga gga gaa agt act cat cag aag caa tac gtt ggt 1178
Met Val Glu Asn Gly Gly Glu Ser Thr His Gln Lys Gln Tyr Val Gly
335 340 345

gga tgt tta cca tca agt ggg ttt gga gat gga ggt aga gaa gct aga 1226
Gly Cys Leu Pro Ser Ser Gly Phe Gly Asp Gly Gly Arg Glu Ala Arg
350 355 360

gtt tcg aga tac aga gag aag agg agg aca agg ttg ttt tct aag aag 1274
Val Ser Arg Tyr Arg Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys
365 370 375

ata cgg tac gag gta cgt aaa ttg aat gca gag aaa aga cca cga atg 1322
Ile Arg Tyr Glu Val Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met
380 385 390 395

```

aaa gga aga ttc gtg aag aga gcc tcg ctc gct gct gct gct tca cca 1370  
 Lys Gly Arg Phe Val Lys Arg Ala Ser Leu Ala Ala Ala Ala Ser Pro  
 400 405 410

tta ggt gtt aat tac tga atagttaata tctattcatg ttatatctca 1418  
 Leu Gly Val Asn Tyr \*  
 415

ctttacaaat ttccgtgaat cttttttctt ctgaaacaac agaagttatt ttggcactta 1478  
 attgtgcttt gaggacttgt atgtacatag aagtaaccaa taataatgtg acttttacta 1538

<210> 172  
 <211> 416  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (18)...(56)  
 <223> Conserved domain

<400> 172  
 Met Lys Ser Leu Ala Asn Ala Val Gly Ala Lys Thr Ala Arg Ala Cys  
 1 5 10 15  
 Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr Cys Ala Ala Asp Asp  
 20 25 30  
 Ala Phe Leu Cys Gln Ser Cys Asp Ser Leu Val His Ser Ala Asn Pro  
 35 40 45  
 Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys Thr Ala Ser Pro Ala  
 50 55 60  
 Val Val Lys His Ser Asn His Ser Ser Ala Ser Pro Pro His Glu Val  
 65 70 75 80  
 Ala Thr Trp His His Gly Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly  
 85 90 95  
 Ser Gly Lys Lys Asn Asn Ser Ser Ile Phe His Asp Leu Val Pro Asp  
 100 105 110  
 Ile Ser Ile Glu Asp Gln Thr Asp Asn Tyr Glu Leu Glu Glu Gln Leu  
 115 120 125  
 Ile Cys Gln Val Pro Val Leu Asp Pro Leu Val Ser Glu Gln Phe Leu  
 130 135 140  
 Asn Asp Val Val Glu Pro Lys Ile Glu Phe Pro Met Ile Arg Ser Gly  
 145 150 155 160  
 Leu Met Ile Glu Glu Glu Glu Asp Asn Ala Glu Ser Cys Leu Asn Gly  
 165 170 175  
 Phe Phe Pro Thr Asp Met Glu Leu Glu Glu Phe Ala Ala Asp Val Glu  
 180 185 190  
 Thr Leu Leu Gly Arg Gly Leu Asp Thr Glu Ser Tyr Ala Met Glu Glu  
 195 200 205  
 Leu Gly Leu Ser Asn Ser Glu Met Phe Lys Ile Glu Lys Asp Glu Ile  
 210 215 220  
 Glu Glu Glu Val Glu Glu Ile Lys Ala Met Ser Met Asp Ile Phe Asp  
 225 230 235 240  
 Asp Asp Arg Lys Asp Val Asp Gly Thr Val Pro Phe Glu Leu Ser Phe  
 245 250 255  
 Asp Tyr Glu Ser Ser His Lys Thr Ser Glu Glu Glu Val Met Lys Asn  
 260 265 270

Val Glu Ser Ser Gly Glu Cys Val Val Lys Val Lys Glu Glu Glu His  
 275 280 285  
 Lys Asn Val Leu Met Leu Arg Leu Asn Tyr Asp Ser Val Ile Ser Thr  
 290 295 300  
 Trp Gly Gly Gln Gly Pro Trp Ser Ser Gly Glu Pro Pro Glu Arg  
 305 310 315 320  
 Asp Met Asp Ile Ser Gly Trp Pro Ala Phe Ser Met Val Glu Asn Gly  
 325 330 335  
 Gly Glu Ser Thr His Gln Lys Gln Tyr Val Gly Gly Cys Leu Pro Ser  
 340 345 350  
 Ser Gly Phe Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg  
 355 360 365  
 Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val  
 370 375 380  
 Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe Val  
 385 390 395 400  
 Lys Arg Ala Ser Leu Ala Ala Ala Ala Ser Pro Leu Gly Val Asn Tyr  
 405 410 415

&lt;210&gt; 173

&lt;211&gt; 1233

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1233)

&lt;400&gt; 173

atg gcc gaa act atg aaa gat att acg atg aag aac gat gaa tca cag 48  
 Met Ala Glu Thr Met Lys Asp Ile Thr Met Lys Asn Asp Glu Ser Gln  
 1 5 10 15  
 gaa gaa gaa atc ccc gat caa ttt ctc tgc tgc gtt tgc ctg gaa ctt 96  
 Glu Glu Glu Ile Pro Asp Gln Phe Leu Cys Cys Val Cys Leu Glu Leu  
 20 25 30  
 ctt tac aag cca att gtg tta tct tgt ggt cat cta tca tgt ttt tgg 144  
 Leu Tyr Lys Pro Ile Val Leu Ser Cys Gly His Leu Ser Cys Phe Trp  
 35 40 45  
 tgc gta cat aag tcc atg aat ggc ttt cgt gag tct cat tgt ccg ata 192  
 Cys Val His Lys Ser Met Asn Gly Phe Arg Glu Ser His Cys Pro Ile  
 50 55 60  
 tgt aga gac ccg tat gtt cac ttt ccc tct gtg tgc cag aag ctt tat 240  
 Cys Arg Asp Pro Tyr Val His Phe Pro Ser Val Cys Gln Lys Leu Tyr  
 65 70 75 80  
 ttt ctg tta aag aag atg tac cca ctt gct cat aag aag aga gaa gaa 288  
 Phe Leu Leu Lys Lys Met Tyr Pro Leu Ala His Lys Lys Arg Glu Glu  
 85 90 95  
 caa gtt tta aag gaa gag caa gaa cga gaa tgt ttt tct cct cag att 336  
 Gln Val Leu Lys Glu Glu Gln Glu Arg Glu Cys Phe Ser Pro Gln Ile  
 100 105 110

gat ctt gtt ttg gat ttg tct gtg tgt agt gga gat tct ctc aat gtc	384
Asp Leu Val Leu Asp Leu Ser Val Cys Ser Gly Asp Ser Leu Asn Val	
115 120 125	
tct gat aaa cag aag gtg gaa gag tgt tcg aat gca gcg aac tta tta	432
Ser Asp Lys Gln Lys Val Glu Glu Cys Ser Asn Ala Ala Asn Leu Leu	
130 135 140	
tct agt tca tca agt aga ggt gac att cca tgt atc ccc aaa aat caa	480
Ser Ser Ser Ser Ser Arg Gly Asp Ile Pro Cys Ile Pro Lys Asn Gln	
145 150 155 160	
gaa ccc aca gat gca aaa gct ctt aat gtt cat gaa aat gaa tta ctt	528
Glu Pro Thr Asp Ala Lys Ala Leu Asn Val His Glu Asn Glu Leu Leu	
165 170 175	
aag gat aac aaa gtc agt aag cag att tcg aaa gat gat ttg ctc tgt	576
Lys Asp Asn Lys Val Ser Lys Gln Ile Ser Lys Asp Asp Leu Leu Cys	
180 185 190	
tca gca tgt aag gag ctg ctt gta cga ccc gta gtt ctc aat tgc gga	624
Ser Ala Cys Lys Glu Leu Leu Val Arg Pro Val Val Leu Asn Cys Gly	
195 200 205	
cat gtg tat tgt gaa gga tgt gta gta gat atg gct gaa gaa agc gaa	672
His Val Tyr Cys Glu Gly Cys Val Val Asp Met Ala Glu Glu Ser Glu	
210 215 220	
aag atc aaa tgt caa gag tgt aat gtt tgt gac cca aga gga ttt cca	720
Lys Ile Lys Cys Gln Glu Cys Asn Val Cys Asp Pro Arg Gly Phe Pro	
225 230 235 240	
aaa gtt tgt ttg att ctt gaa cag ctt ttg gag gaa aac ttt cct gaa	768
Lys Val Cys Leu Ile Leu Glu Gln Leu Leu Glu Glu Asn Phe Pro Glu	
245 250 255	
gaa tac aat tca aga agc agt aag gtt cag aaa acg ctc gcc cat aat	816
Glu Tyr Asn Ser Arg Ser Ser Lys Val Gln Lys Thr Leu Ala His Asn	
260 265 270	
agc aaa gga aat att caa agc tat ctc aaa gaa ggc ccg tcc tta tca	864
Ser Lys Gly Asn Ile Gln Ser Tyr Leu Lys Glu Gly Pro Ser Leu Ser	
275 280 285	
aac gac aat aac aat gat gat ccc tgg ttg gca aac cct gga tca aat	912
Asn Asp Asn Asn Asp Asp Pro Trp Leu Ala Asn Pro Gly Ser Asn	
290 295 300	
gtt cac ttt gga gct ggt tgt gat tct tgt ggg gtg tat cca atc ata	960
Val His Phe Gly Ala Gly Cys Asp Ser Cys Gly Val Tyr Pro Ile Ile	
305 310 315 320	
ggg gat cga tac aga tgc aaa gac tgc aag gag gaa att ggg tat gac	1008
Gly Asp Arg Tyr Arg Cys Lys Asp Cys Lys Glu Glu Ile Gly Tyr Asp	
325 330 335	

```

ctt tgc aaa gac tgt tac gag act cct tcg aaa gtt cca ggg aga ttc 1056
Leu Cys Lys Asp Cys Tyr Glu Thr Pro Ser Lys Val Pro Gly Arg Phe
      340                      345                      350

aac cag caa cac act cct gac cac agg ctt gag ctt gca cgg tct cct 1104
Asn Gln Gln His Thr Pro Asp His Arg Leu Glu Leu Ala Arg Ser Pro
      355                      360                      365

cag gtt ctg atc aat ttc aat tct atc ggt atc ctt ctc gga ccc gtt 1152
Gln Val Leu Ile Asn Phe Asn Ser Ile Gly Ile Leu Leu Gly Pro Val
      370                      375                      380

atc tca aat gaa ggc atg gat aca gat gaa ggc gag gaa ggg cct cct 1200
Ile Ser Asn Glu Gly Met Asp Thr Asp Glu Gly Glu Glu Gly Pro Pro
      385                      390                      395                      400

ggt tct tct aat gag tca tca agc aca gaa tga 1233
Gly Ser Ser Asn Glu Ser Ser Ser Thr Glu *
      405                      410

```

<210> 174  
 <211> 410  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (20)...(50)  
 <223> Conserved domain

```

<400> 174
Met Ala Glu Thr Met Lys Asp Ile Thr Met Lys Asn Asp Glu Ser Gln
 1      5      10      15
Glu Glu Glu Ile Pro Asp Gln Phe Leu Cys Cys Val Cys Leu Glu Leu
      20      25      30
Leu Tyr Lys Pro Ile Val Leu Ser Cys Gly His Leu Ser Cys Phe Trp
      35      40      45
Cys Val His Lys Ser Met Asn Gly Phe Arg Glu Ser His Cys Pro Ile
      50      55      60
Cys Arg Asp Pro Tyr Val His Phe Pro Ser Val Cys Gln Lys Leu Tyr
      65      70      75      80
Phe Leu Leu Lys Lys Met Tyr Pro Leu Ala His Lys Lys Arg Glu Glu
      85      90      95
Gln Val Leu Lys Glu Glu Gln Glu Arg Glu Cys Phe Ser Pro Gln Ile
      100     105     110
Asp Leu Val Leu Asp Leu Ser Val Cys Ser Gly Asp Ser Leu Asn Val
      115     120     125
Ser Asp Lys Gln Lys Val Glu Glu Cys Ser Asn Ala Ala Asn Leu Leu
      130     135     140
Ser Ser Ser Ser Ser Arg Gly Asp Ile Pro Cys Ile Pro Lys Asn Gln
      145     150     155     160
Glu Pro Thr Asp Ala Lys Ala Leu Asn Val His Glu Asn Glu Leu Leu
      165     170     175
Lys Asp Asn Lys Val Ser Lys Gln Ile Ser Lys Asp Asp Leu Leu Cys
      180     185     190
Ser Ala Cys Lys Glu Leu Leu Val Arg Pro Val Val Leu Asn Cys Gly

```

195	200	205
His Val Tyr Cys Glu Gly	Cys Val Val Asp Met	Ala Glu Glu Ser Glu
210	215	220
Lys Ile Lys Cys Gln Glu	Cys Asn Val Cys Asp	Pro Arg Gly Phe Pro
225	230	235
Lys Val Cys Leu Ile Leu	Glu Gln Leu Leu Glu	Glu Asn Phe Pro Glu
245	250	255
Glu Tyr Asn Ser Arg Ser	Ser Lys Val Gln Lys Thr	Leu Ala His Asn
260	265	270
Ser Lys Gly Asn Ile Gln	Ser Tyr Leu Lys Glu Gly	Pro Ser Leu Ser
275	280	285
Asn Asp Asn Asn Asn Asp	Asp Pro Trp Leu Ala Asn	Pro Gly Ser Asn
290	295	300
Val His Phe Gly Ala Gly	Cys Asp Ser Cys Gly Val	Tyr Pro Ile Ile
305	310	315
Gly Asp Arg Tyr Arg Cys	Lys Asp Cys Lys Glu Glu	Ile Gly Tyr Asp
325	330	335
Leu Cys Lys Asp Cys Tyr	Glu Thr Pro Ser Lys Val	Pro Gly Arg Phe
340	345	350
Asn Gln Gln His Thr Pro	Asp His Arg Leu Glu Leu	Ala Arg Ser Pro
355	360	365
Gln Val Leu Ile Asn Phe	Asn Ser Ile Gly Ile Leu	Leu Gly Pro Val
370	375	380
Ile Ser Asn Glu Gly Met	Asp Thr Asp Glu Gly Glu	Glu Gly Pro Pro
385	390	395
Gly Ser Ser Asn Glu Ser	Ser Ser Thr Glu	
405	410	

&lt;210&gt; 175

&lt;211&gt; 1031

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (71)...(1003)

&lt;400&gt; 175

atatgtactt gtacaccaac ccaccaaag agataaaaga ggaacaaaa actcgaaaag	60
agagagatat atg ggt gag gtg gct tat atg gac gaa gga gac cta gaa	109
Met Gly Glu Val Ala Tyr Met Asp Glu Gly Asp Leu Glu	
1 5 10	
gca ata gtc aga ggc tac tcc ggc tcc gga gac gcg ttt tcc ggc gaa	157
Ala Ile Val Arg Gly Tyr Ser Gly Ser Gly Asp Ala Phe Ser Gly Glu	
15 20 25	
agt tcc ggt acg ttt tca cct tcg ttt tgc cta ccg atg gag acg tct	205
Ser Ser Gly Thr Phe Ser Pro Ser Phe Cys Leu Pro Met Glu Thr Ser	
30 35 40 45	
agt ttc tac gaa ccg gag atg gag aca agt ggc tta gat gag ctc ggt	253
Ser Phe Tyr Glu Pro Glu Met Glu Thr Ser Gly Leu Asp Glu Leu Gly	
50 55 60	
gaa ctt tac aaa ccc ttt tac cct ttc tcc aca caa acg atc ctc aca	301



Glu Leu Tyr Lys Pro Phe Tyr Pro Phe Ser Thr Gln Thr Ile Leu Thr	
65 70 75	
agc tcg gtc tct ctc cct gaa gat tca aaa cct ttc cga gat gac aag	349
Ser Ser Val Ser Leu Pro Glu Asp Ser Lys Pro Phe Arg Asp Asp Lys	
80 85 90	
aaa caa cga tca cat ggt tgt ctt tta tcc aac gga tca aga gct gat	397
Lys Gln Arg Ser His Gly Cys Leu Leu Ser Asn Gly Ser Arg Ala Asp	
95 100 105	
cat atc cga att tca gaa tcc aaa tca aag aaa agc aag aag aat caa	445
His Ile Arg Ile Ser Glu Ser Lys Ser Lys Lys Ser Lys Lys Asn Gln	
110 115 120 125	
cag aag aga gtt gtt gag caa gtg aaa gaa gag aat ctg ttg tcg gac	493
Gln Lys Arg Val Val Glu Gln Val Lys Glu Glu Asn Leu Leu Ser Asp	
130 135 140	
gca tgg gcg tgg cgt aaa tac ggg cag aaa ccc atc aaa gga tct cca	541
Ala Trp Ala Trp Arg Lys Tyr Gly Gln Lys Pro Ile Lys Gly Ser Pro	
145 150 155	
tac cca agg agt tat tac aga tgc agt agc tca aaa ggg tgt ttg gca	589
Tyr Pro Arg Ser Tyr Tyr Arg Cys Ser Ser Ser Lys Gly Cys Leu Ala	
160 165 170	
aga aaa caa gtc gaa aga aat cct caa aac ccg gag aaa ttc acc ata	637
Arg Lys Gln Val Glu Arg Asn Pro Gln Asn Pro Glu Lys Phe Thr Ile	
175 180 185	
aca tac act aat gag cac aat cat gaa cta cca acc cgg aga aac tca	685
Thr Tyr Thr Asn Glu His Asn His Glu Leu Pro Thr Arg Arg Asn Ser	
190 195 200 205	
tta gcc ggt tcg act cga gca aaa act tcc caa ccc aaa cca acc tta	733
Leu Ala Gly Ser Thr Arg Ala Lys Thr Ser Gln Pro Lys Pro Thr Leu	
210 215 220	
acc aaa aaa tcc gaa aaa gaa gtt gtt tct tcc cct aca agt aat cct	781
Thr Lys Lys Ser Glu Lys Glu Val Val Ser Ser Pro Thr Ser Asn Pro	
225 230 235	
atg atc cca tcc gct gat gaa tct tct gtt gcg gtt caa gaa atg agc	829
Met Ile Pro Ser Ala Asp Glu Ser Ser Val Ala Val Gln Glu Met Ser	
240 245 250	
gtt gcg gaa acg agt acg cac caa gcg gct gga gca atc gag ggc cgc	877
Val Ala Glu Thr Ser Thr His Gln Ala Ala Gly Ala Ile Glu Gly Arg	
255 260 265	
cgc ttg agt aac ggt tta cca tcg gat ttg atg tcc ggg agc gga act	925
Arg Leu Ser Asn Gly Leu Pro Ser Asp Leu Met Ser Gly Ser Gly Thr	
270 275 280 285	
ttt cca agt ttt acc ggt gac ttc gat gaa cta ttg aat agc caa gag	973
Phe Pro Ser Phe Thr Gly Asp Phe Asp Glu Leu Leu Asn Ser Gln Glu	

290 295 300  
 ttc ttc agt ggg tat tta tgg aat tac tag agagcattag gtgtatgtat 1023  
 Phe Phe Ser Gly Tyr Leu Trp Asn Tyr \*  
 305 310  
 atatatat 1031  
 <210> 176  
 <211> 310  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 176  
 Met Gly Glu Val Ala Tyr Met Asp Glu Gly Asp Leu Glu Ala Ile Val  
 1 5 10 15  
 Arg Gly Tyr Ser Gly Ser Gly Asp Ala Phe Ser Gly Glu Ser Ser Gly  
 20 25 30  
 Thr Phe Ser Pro Ser Phe Cys Leu Pro Met Glu Thr Ser Ser Phe Tyr  
 35 40 45  
 Glu Pro Glu Met Glu Thr Ser Gly Leu Asp Glu Leu Gly Glu Leu Tyr  
 50 55 60  
 Lys Pro Phe Tyr Pro Phe Ser Thr Gln Thr Ile Leu Thr Ser Ser Val  
 65 70 75 80  
 Ser Leu Pro Glu Asp Ser Lys Pro Phe Arg Asp Asp Lys Lys Gln Arg  
 85 90 95  
 Ser His Gly Cys Leu Leu Ser Asn Gly Ser Arg Ala Asp His Ile Arg  
 100 105 110  
 Ile Ser Glu Ser Lys Ser Lys Lys Ser Lys Lys Asn Gln Gln Lys Arg  
 115 120 125  
 Val Val Glu Gln Val Lys Glu Glu Asn Leu Leu Ser Asp Ala Trp Ala  
 130 135 140  
 Trp Arg Lys Tyr Gly Gln Lys Pro Ile Lys Gly Ser Pro Tyr Pro Arg  
 145 150 155 160  
 Ser Tyr Tyr Arg Cys Ser Ser Ser Lys Gly Cys Leu Ala Arg Lys Gln  
 165 170 175  
 Val Glu Arg Asn Pro Gln Asn Pro Glu Lys Phe Thr Ile Thr Tyr Thr  
 180 185 190  
 Asn Glu His Asn His Glu Leu Pro Thr Arg Arg Asn Ser Leu Ala Gly  
 195 200 205  
 Ser Thr Arg Ala Lys Thr Ser Gln Pro Lys Pro Thr Leu Thr Lys Lys  
 210 215 220  
 Ser Glu Lys Glu Val Val Ser Ser Pro Thr Ser Asn Pro Met Ile Pro  
 225 230 235 240  
 Ser Ala Asp Glu Ser Ser Val Ala Val Gln Glu Met Ser Val Ala Glu  
 245 250 255  
 Thr Ser Thr His Gln Ala Ala Gly Ala Ile Glu Gly Arg Arg Leu Ser  
 260 265 270  
 Asn Gly Leu Pro Ser Asp Leu Met Ser Gly Ser Gly Thr Phe Pro Ser  
 275 280 285  
 Phe Thr Gly Asp Phe Asp Glu Leu Leu Asn Ser Gln Glu Phe Phe Ser  
 290 295 300  
 Gly Tyr Leu Trp Asn Tyr  
 305 310

&lt;210&gt; 177

&lt;211&gt; 1320

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1320)

&lt;400&gt; 177

```

atg ttg tcc aac aag aac aca aac aca tgt tgt gtg gtt tct tct tct 48
Met Leu Ser Asn Lys Asn Thr Asn Thr Cys Cys Val Val Ser Ser Ser
1 5 10 15

tct tct gat cct ttc ctc tcc tct tca gaa aat ggg gtc acc acc aca 96
Ser Ser Asp Pro Phe Leu Ser Ser Ser Glu Asn Gly Val Thr Thr Thr
20 25 30

aac aca tcc act cag aag agg aaa aga aga cct gca ggt acc cca gat 144
Asn Thr Ser Thr Gln Lys Arg Lys Arg Arg Pro Ala Gly Thr Pro Asp
35 40 45

cca gat gca gaa gtt gtg tct tta tca cca aga act ctt ctt gaa tca 192
Pro Asp Ala Glu Val Val Ser Leu Ser Pro Arg Thr Leu Leu Glu Ser
50 55 60

gac aga tac ata tgt gag atc tgt aac caa ggg ttt caa aga gac cag 240
Asp Arg Tyr Ile Cys Glu Ile Cys Asn Gln Gly Phe Gln Arg Asp Gln
65 70 75 80

aat ctc cag atg cat aga aga cgt cac aaa gtt cca tgg aag ctt ctt 288
Asn Leu Gln Met His Arg Arg Arg His Lys Val Pro Trp Lys Leu Leu
85 90 95

aaa aga gac aac aac ata gag gtg aag aaa cga gtc tat gtt tgc cct 336
Lys Arg Asp Asn Asn Ile Glu Val Lys Lys Arg Val Tyr Val Cys Pro
100 105 110

gaa ccc act tgc ctt cac cat aat cct tgt cat gct ctg gga gat ctt 384
Glu Pro Thr Cys Leu His His Asn Pro Cys His Ala Leu Gly Asp Leu
115 120 125

gtc ggc atc aaa aaa cat ttc aga aga aaa cat agt aac cat aag caa 432
Val Gly Ile Lys Lys His Phe Arg Arg Lys His Ser Asn His Lys Gln
130 135 140

tgg gtt tgt gag aga tgc tct aaa ggt tat gct gtt caa tct gat tac 480
Trp Val Cys Glu Arg Cys Ser Lys Gly Tyr Ala Val Gln Ser Asp Tyr
145 150 155 160

aaa gct cat ctc aaa act tgt ggt act aga gga cat tct tgt gac tgt 528
Lys Ala His Leu Lys Thr Cys Gly Thr Arg Gly His Ser Cys Asp Cys
165 170 175

ggg cgt gtc ttc tcc agg gtg gag agt ttt att gaa cat caa gat aat 576
Gly Arg Val Phe Ser Arg Val Glu Ser Phe Ile Glu His Gln Asp Asn
180 185 190

```

tgt tcc gca cgg aga gtt cac cgt gaa ccg cct cga cca cca caa acc	624
Cys Ser Ala Arg Arg Val His Arg Glu Pro Pro Arg Pro Pro Gln Thr	
195 200 205	
gcg gta aca gtc ccg gca tgc tct tct aga acc gcc tca acc gtc agc	672
Ala Val Thr Val Pro Ala Cys Ser Ser Arg Thr Ala Ser Thr Val Ser	
210 215 220	
act ccg tct agt gaa acc aat tac ggc ggt aca gtt gcg gtt acg act	720
Thr Pro Ser Ser Glu Thr Asn Tyr Gly Gly Thr Val Ala Val Thr Thr	
225 230 235 240	
cct caa cct cta gaa ggc cgt cca att cat cag aga atc tca tct tca	768
Pro Gln Pro Leu Glu Gly Arg Pro Ile His Gln Arg Ile Ser Ser Ser	
245 250 255	
att ctc acc aac tca tca aac aat ctc aac ctc gaa ctc caa ctt ctt	816
Ile Leu Thr Asn Ser Ser Asn Asn Leu Asn Leu Glu Leu Gln Leu Leu	
260 265 270	
cca tta tcg tcg aat caa aac cct aat caa gaa aac caa caa caa aaa	864
Pro Leu Ser Ser Asn Gln Asn Pro Asn Gln Glu Asn Gln Gln Gln Lys	
275 280 285	
gtt aaa gaa cca tct cat cat cat aat cat aat cat gat acc aca aac	912
Val Lys Glu Pro Ser His His His Asn His Asn His Asp Thr Thr Asn	
290 295 300	
tta aac ctc tcc att gca cca tca tca tca tat caa cat tac aac aac	960
Leu Asn Leu Ser Ile Ala Pro Ser Ser Ser Tyr Gln His Tyr Asn Asn	
305 310 315 320	
ttt gat cgt ata aaa gag att atg gcg agc gag caa att atg aag ata	1008
Phe Asp Arg Ile Lys Glu Ile Met Ala Ser Glu Gln Ile Met Lys Ile	
325 330 335	
gcg atg aag gag aaa gct tac gct gag gaa gct aaa aga gaa gcc aag	1056
Ala Met Lys Glu Lys Ala Tyr Ala Glu Glu Ala Lys Arg Glu Ala Lys	
340 345 350	
agg caa cga gag ata gcg gaa aac gag ttt gca aat gcc aaa aag att	1104
Arg Gln Arg Glu Ile Ala Glu Asn Glu Phe Ala Asn Ala Lys Lys Ile	
355 360 365	
agg caa aaa gca caa gct gaa ctc gag aga gct aag ttt tta aag gaa	1152
Arg Gln Lys Ala Gln Ala Glu Leu Glu Arg Ala Lys Phe Leu Lys Glu	
370 375 380	
caa tct atg aag aag ata agt tca acg atc atg cag gtc act tgt caa	1200
Gln Ser Met Lys Lys Ile Ser Ser Thr Ile Met Gln Val Thr Cys Gln	
385 390 395 400	
aca tgt aaa gga cag ttt caa gca gtt gcg gtt ccg gcg gct acg gct	1248
Thr Cys Lys Gly Gln Phe Gln Ala Val Ala Val Pro Ala Ala Thr Ala	
405 410 415	
gat gag acg tct ctt gtc gtg agt tac atg tcg tca gcg aat act gac	1296

Asp Glu Thr Ser Leu Val Val Ser Tyr Met Ser Ser Ala Asn Thr Asp  
 420 425 430

gga gag ttg gaa aat ggt ttt taa  
 Gly Glu Leu Glu Asn Gly Phe \*  
 435

1320

<210> 178  
 <211> 439  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (69)...(89)  
 <223> Conserved domain

<221> DOMAIN  
 <222> (111)...(139)  
 <223> Conserved domain

<400> 178  
 Met Leu Ser Asn Lys Asn Thr Asn Thr Cys Cys Val Val Ser Ser Ser  
 1 5 10 15  
 Ser Ser Asp Pro Phe Leu Ser Ser Ser Glu Asn Gly Val Thr Thr Thr  
 20 25 30  
 Asn Thr Ser Thr Gln Lys Arg Lys Arg Arg Pro Ala Gly Thr Pro Asp  
 35 40 45  
 Pro Asp Ala Glu Val Val Ser Leu Ser Pro Arg Thr Leu Leu Glu Ser  
 50 55 60  
 Asp Arg Tyr Ile Cys Glu Ile Cys Asn Gln Gly Phe Gln Arg Asp Gln  
 65 70 75 80  
 Asn Leu Gln Met His Arg Arg Arg His Lys Val Pro Trp Lys Leu Leu  
 85 90 95  
 Lys Arg Asp Asn Asn Ile Glu Val Lys Lys Arg Val Tyr Val Cys Pro  
 100 105 110  
 Glu Pro Thr Cys Leu His His Asn Pro Cys His Ala Leu Gly Asp Leu  
 115 120 125  
 Val Gly Ile Lys Lys His Phe Arg Arg Lys His Ser Asn His Lys Gln  
 130 135 140  
 Trp Val Cys Glu Arg Cys Ser Lys Gly Tyr Ala Val Gln Ser Asp Tyr  
 145 150 155 160  
 Lys Ala His Leu Lys Thr Cys Gly Thr Arg Gly His Ser Cys Asp Cys  
 165 170 175  
 Gly Arg Val Phe Ser Arg Val Glu Ser Phe Ile Glu His Gln Asp Asn  
 180 185 190  
 Cys Ser Ala Arg Arg Val His Arg Glu Pro Pro Arg Pro Pro Gln Thr  
 195 200 205  
 Ala Val Thr Val Pro Ala Cys Ser Ser Arg Thr Ala Ser Thr Val Ser  
 210 215 220  
 Thr Pro Ser Ser Glu Thr Asn Tyr Gly Gly Thr Val Ala Val Thr Thr  
 225 230 235 240  
 Pro Gln Pro Leu Glu Gly Arg Pro Ile His Gln Arg Ile Ser Ser Ser  
 245 250 255  
 Ile Leu Thr Asn Ser Ser Asn Asn Leu Asn Leu Glu Leu Gln Leu Leu  
 260 265 270

Pro Leu Ser Ser Asn Gln Asn Pro Asn Gln Glu Asn Gln Gln Gln Lys  
           275                          280                          285  
 Val Lys Glu Pro Ser His His His Asn His Asn His Asp Thr Thr Asn  
           290                          295                          300  
 Leu Asn Leu Ser Ile Ala Pro Ser Ser Ser Tyr Gln His Tyr Asn Asn  
 305                          310                          315                          320  
 Phe Asp Arg Ile Lys Glu Ile Met Ala Ser Glu Gln Ile Met Lys Ile  
                           325                          330                          335  
 Ala Met Lys Glu Lys Ala Tyr Ala Glu Glu Ala Lys Arg Glu Ala Lys  
                           340                          345                          350  
 Arg Gln Arg Glu Ile Ala Glu Asn Glu Phe Ala Asn Ala Lys Lys Ile  
                           355                          360                          365  
 Arg Gln Lys Ala Gln Ala Glu Leu Glu Arg Ala Lys Phe Leu Lys Glu  
 370                          375                          380  
 Gln Ser Met Lys Lys Ile Ser Ser Thr Ile Met Gln Val Thr Cys Gln  
 385                          390                          395                          400  
 Thr Cys Lys Gly Gln Phe Gln Ala Val Ala Val Pro Ala Ala Thr Ala  
                           405                          410                          415  
 Asp Glu Thr Ser Leu Val Val Ser Tyr Met Ser Ser Ala Asn Thr Asp  
                           420                          425                          430  
 Gly Glu Leu Glu Asn Gly Phe  
                           435

&lt;210&gt; 179

&lt;211&gt; 1521

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1521)

&lt;400&gt; 179

atg gag tac gat gga cgg aga ccc tac tct gtt cta acg aga aat gaa 48  
 Met Glu Tyr Asp Gly Arg Arg Pro Tyr Ser Val Leu Thr Arg Asn Glu  
   1                          5                          10                          15

att acc gta aag atg aag aaa cag atc aat gaa atc tct gac atc ttc 96  
 Ile Thr Val Lys Met Lys Lys Gln Ile Asn Glu Ile Ser Asp Ile Phe  
                           20                          25                          30

ttc ata tca aac tcc gat gcg acc gta ctt ctt atg tat ctc cga tgg 144  
 Phe Ile Ser Asn Ser Asp Ala Thr Val Leu Leu Met Tyr Leu Arg Trp  
                           35                          40                          45

gac tcg ctt cgt gtt tca gag cgt tta ggt gaa aac aag gag aaa tta 192  
 Asp Ser Leu Arg Val Ser Glu Arg Leu Gly Glu Asn Lys Glu Lys Leu  
                           50                          55                          60

ttg atg gat tca ggt ttg aaa tca gtc atg atc gat ccc agt cca gat 240  
 Leu Met Asp Ser Gly Leu Lys Ser Val Met Ile Asp Pro Ser Pro Asp  
                           65                          70                          75                          80

tca tct agc gag att agc ttg gag act gat gtt tat gag ttt gat ggt 288  
 Ser Ser Ser Glu Ile Ser Leu Glu Thr Asp Val Tyr Glu Phe Asp Gly  
                           85                          90                          95

gat aat gat tta atc tcc atg ccc ttt tgt tct cac aag ttc gac tca	336
Asp Asn Asp Leu Ile Ser Met Pro Phe Cys Ser His Lys Phe Asp Ser	
100 105 110	
aaa tat tgg aga gag tat ctt gag aag aat ttc tac tat gtg gag aaa	384
Lys Tyr Trp Arg Glu Tyr Leu Glu Lys Asn Phe Tyr Tyr Val Glu Lys	
115 120 125	
atc caa acc aca atc tct tgt cct gat caa gat tgt cga tct gcg gtt	432
Ile Gln Thr Thr Ile Ser Cys Pro Asp Gln Asp Cys Arg Ser Ala Val	
130 135 140	
gga cca gac acg att gag aaa cta aca gta cgt gat caa gag atg tac	480
Gly Pro Asp Thr Ile Glu Lys Leu Thr Val Arg Asp Gln Glu Met Tyr	
145 150 155 160	
gag agg tac att tgg aga tct tac att gaa gga aac aag gta ttg atg	528
Glu Arg Tyr Ile Trp Arg Ser Tyr Ile Glu Gly Asn Lys Val Leu Met	
165 170 175	
atc aaa cag tgt cct gca cgg aat tgc gat tat gtt att gag ttt cat	576
Ile Lys Gln Cys Pro Ala Arg Asn Cys Asp Tyr Val Ile Glu Phe His	
180 185 190	
caa gag aat gat gat gat gat gag tat agt tta aac gtg gtg tgt atc	624
Gln Glu Asn Asp Asp Asp Asp Glu Tyr Ser Leu Asn Val Val Cys Ile	
195 200 205	
tgt ggt cat atc ttc tgt tgg aga tgc aga ctt gag tca cac aga cca	672
Cys Gly His Ile Phe Cys Trp Arg Cys Arg Leu Glu Ser His Arg Pro	
210 215 220	
gtg agt tgt aac aag gcc tct gat tgg ttg tgt agt gcc aca atg aaa	720
Val Ser Cys Asn Lys Ala Ser Asp Trp Leu Cys Ser Ala Thr Met Lys	
225 230 235 240	
ata tca gat gaa tca ttt agc ctt tat ccg acc aaa acc aag aca gtg	768
Ile Ser Asp Glu Ser Phe Ser Leu Tyr Pro Thr Lys Thr Lys Thr Val	
245 250 255	
act tgc cct cat tgc ctc tgt tct ctg gag tct gat aca aaa atg ccc	816
Thr Cys Pro His Cys Leu Cys Ser Leu Glu Ser Asp Thr Lys Met Pro	
260 265 270	
cag ttc ttg act tgt gtc tgc agg ctt cgt ttc tgt tcg agg tgt ttg	864
Gln Phe Leu Thr Cys Val Cys Arg Leu Arg Phe Cys Ser Arg Cys Leu	
275 280 285	
cgg tca gag gaa gcc cac aaa att gaa gca gta gac tct gga ttc tgt	912
Arg Ser Glu Glu Ala His Lys Ile Glu Ala Val Asp Ser Gly Phe Cys	
290 295 300	
att aaa aca gaa gtg ggg att tta tgt gag gat cgc tgg aac gtg tgt	960
Ile Lys Thr Glu Val Gly Ile Leu Cys Glu Asp Arg Trp Asn Val Cys	
305 310 315 320	

```

cag aag tta ctg gag caa gcc aaa tcc gat cta gaa gct ttc gag gaa 1008
Gln Lys Leu Leu Glu Gln Ala Lys Ser Asp Leu Glu Ala Phe Glu Glu
          325                330                335

acc aac atc aag aaa ccg agt gac tta tta agg gag caa gac att atg 1056
Thr Asn Ile Lys Lys Pro Ser Asp Leu Leu Arg Glu Gln Asp Ile Met
          340                345                350

att ata aga gaa ggg ttg atg ttg att gtt caa tgc aga cga gtc ctt 1104
Ile Ile Arg Glu Gly Leu Met Leu Ile Val Gln Cys Arg Arg Val Leu
          355                360                365

aaa tgg tgt tgt gtg tat gac tat ttc cac aca gag tat gag aac tcg 1152
Lys Trp Cys Cys Val Tyr Asp Tyr Phe His Thr Glu Tyr Glu Asn Ser
          370                375                380

aag gag tat cta cga tac ctc caa gga aat gca atc gcc acg ctc cag 1200
Lys Glu Tyr Leu Arg Tyr Leu Gln Gly Asn Ala Ile Ala Thr Leu Gln
          385                390                395                400

agt tac tca aac act tta caa gag caa aaa gat atc gtc ctc gct gca 1248
Ser Tyr Ser Asn Thr Leu Gln Glu Gln Lys Asp Ile Val Leu Ala Ala
          405                410                415

gct act tat gaa gag tgt act ttc ttc aga cat acg ata cct act gcg 1296
Ala Thr Tyr Glu Glu Cys Thr Phe Phe Arg His Thr Ile Pro Thr Ala
          420                425                430

aca agc aac att ggt aac tac ttt tat gat ttt atg aaa act tta caa 1344
Thr Ser Asn Ile Gly Asn Tyr Phe Tyr Asp Phe Met Lys Thr Leu Gln
          435                440                445

gat ggt ttg gtc gat gta aag gtg aag tca tac aac ggc ggg act ggt 1392
Asp Gly Leu Val Asp Val Lys Val Lys Ser Tyr Asn Gly Gly Thr Gly
          450                455                460

cct ttc tgg tac tgt gat cgc tgc acg tat gca aac act tgg gag gat 1440
Pro Phe Trp Tyr Cys Asp Arg Cys Thr Tyr Ala Asn Thr Trp Glu Asp
          465                470                475                480

aat gag tgt gag atg tgc tat gat gat agt gcc tct ctt gtg gga gag 1488
Asn Glu Cys Glu Met Cys Tyr Asp Asp Ser Ala Ser Leu Val Gly Glu
          485                490                495

ata agt gat ttg ttt ctt aac aaa gtt tct taa 1521
Ile Ser Asp Leu Phe Leu Asn Lys Val Ser *
          500                505

```

&lt;210&gt; 180

&lt;211&gt; 506

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (214)...(287)



&lt;223&gt; Conserved domain

&lt;400&gt; 180

```

Met Glu Tyr Asp Gly Arg Arg Pro Tyr Ser Val Leu Thr Arg Asn Glu
 1          5          10          15
Ile Thr Val Lys Met Lys Lys Gln Ile Asn Glu Ile Ser Asp Ile Phe
          20          25          30
Phe Ile Ser Asn Ser Asp Ala Thr Val Leu Leu Met Tyr Leu Arg Trp
          35          40          45
Asp Ser Leu Arg Val Ser Glu Arg Leu Gly Glu Asn Lys Glu Lys Leu
          50          55          60
Leu Met Asp Ser Gly Leu Lys Ser Val Met Ile Asp Pro Ser Pro Asp
          65          70          75          80
Ser Ser Ser Glu Ile Ser Leu Glu Thr Asp Val Tyr Glu Phe Asp Gly
          85          90          95
Asp Asn Asp Leu Ile Ser Met Pro Phe Cys Ser His Lys Phe Asp Ser
          100          105          110
Lys Tyr Trp Arg Glu Tyr Leu Glu Lys Asn Phe Tyr Tyr Val Glu Lys
          115          120          125
Ile Gln Thr Thr Ile Ser Cys Pro Asp Gln Asp Cys Arg Ser Ala Val
          130          135          140
Gly Pro Asp Thr Ile Glu Lys Leu Thr Val Arg Asp Gln Glu Met Tyr
          145          150          155          160
Glu Arg Tyr Ile Trp Arg Ser Tyr Ile Glu Gly Asn Lys Val Leu Met
          165          170          175
Ile Lys Gln Cys Pro Ala Arg Asn Cys Asp Tyr Val Ile Glu Phe His
          180          185          190
Gln Glu Asn Asp Asp Asp Asp Glu Tyr Ser Leu Asn Val Val Cys Ile
          195          200          205
Cys Gly His Ile Phe Cys Trp Arg Cys Arg Leu Glu Ser His Arg Pro
          210          215          220
Val Ser Cys Asn Lys Ala Ser Asp Trp Leu Cys Ser Ala Thr Met Lys
          225          230          235          240
Ile Ser Asp Glu Ser Phe Ser Leu Tyr Pro Thr Lys Thr Lys Thr Val
          245          250          255
Thr Cys Pro His Cys Leu Cys Ser Leu Glu Ser Asp Thr Lys Met Pro
          260          265          270
Gln Phe Leu Thr Cys Val Cys Arg Leu Arg Phe Cys Ser Arg Cys Leu
          275          280          285
Arg Ser Glu Glu Ala His Lys Ile Glu Ala Val Asp Ser Gly Phe Cys
          290          295          300
Ile Lys Thr Glu Val Gly Ile Leu Cys Glu Asp Arg Trp Asn Val Cys
          305          310          315          320
Gln Lys Leu Leu Glu Gln Ala Lys Ser Asp Leu Glu Ala Phe Glu Glu
          325          330          335
Thr Asn Ile Lys Lys Pro Ser Asp Leu Leu Arg Glu Gln Asp Ile Met
          340          345          350
Ile Ile Arg Glu Gly Leu Met Leu Ile Val Gln Cys Arg Arg Val Leu
          355          360          365
Lys Trp Cys Cys Val Tyr Asp Tyr Phe His Thr Glu Tyr Glu Asn Ser
          370          375          380
Lys Glu Tyr Leu Arg Tyr Leu Gln Gly Asn Ala Ile Ala Thr Leu Gln
          385          390          395          400
Ser Tyr Ser Asn Thr Leu Gln Glu Gln Lys Asp Ile Val Leu Ala Ala
          405          410          415
Ala Thr Tyr Glu Glu Cys Thr Phe Phe Arg His Thr Ile Pro Thr Ala
          420          425          430

```

```

Thr Ser Asn Ile Gly Asn Tyr Phe Tyr Asp Phe Met Lys Thr Leu Gln
      435                440                445
Asp Gly Leu Val Asp Val Lys Val Lys Ser Tyr Asn Gly Gly Thr Gly
      450                455                460
Pro Phe Trp Tyr Cys Asp Arg Cys Thr Tyr Ala Asn Thr Trp Glu Asp
      465                470                475                480
Asn Glu Cys Glu Met Cys Tyr Asp Asp Ser Ala Ser Leu Val Gly Glu
      485                490                495
Ile Ser Asp Leu Phe Leu Asn Lys Val Ser
      500                505

```

<210> 181  
 <211> 955  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (54)...(785)

```

<400> 181
tcgacccacc ttgaatctga gagagagaga gaagtacgtc cccacacctt gac atg      56
                                         Met
                                         1

gtc atg gag ccc aag aag aac caa aat cta cca agt ttc tta aac cca      104
Val Met Glu Pro Lys Lys Asn Gln Asn Leu Pro Ser Phe Leu Asn Pro
      5                10                15

tca cga cag aat cag gac aac gac aag aag agg aaa caa aca gag gtt      152
Ser Arg Gln Asn Gln Asp Asn Asp Lys Lys Arg Lys Gln Thr Glu Val
      20                25                30

aaa ggt ttc gac att gtg gtc ggc gaa aag agg aag aag aag gag aat      200
Lys Gly Phe Asp Ile Val Val Gly Glu Lys Arg Lys Lys Lys Glu Asn
      35                40                45

gaa gag gaa gac caa gaa att cag att ctt tat gag aag gag aag aag      248
Glu Glu Glu Asp Gln Glu Ile Gln Ile Leu Tyr Glu Lys Glu Lys Lys
      50                55                60                65

aaa cca aac aaa gat cgt cac ctt aaa gtt gaa gga aga ggt cgt aga      296
Lys Pro Asn Lys Asp Arg His Leu Lys Val Glu Gly Arg Gly Arg Arg
      70                75                80

gtt agg tta cct cca ctc tgt gca gca agg att tat caa ttg act aaa      344
Val Arg Leu Pro Pro Leu Cys Ala Ala Arg Ile Tyr Gln Leu Thr Lys
      85                90                95

gaa tta ggt cac aaa tca gat ggt gag act ctt gaa tgg ttg ctt caa      392
Glu Leu Gly His Lys Ser Asp Gly Glu Thr Leu Glu Trp Leu Leu Gln
      100                105                110

cat gct gag cca tcg ata ctc tct gct act gta aat ggt atc aaa ccc      440
His Ala Glu Pro Ser Ile Leu Ser Ala Thr Val Asn Gly Ile Lys Pro
      115                120                125

```

```

act gag tct gtt gtt tct caa cct cct ctc acg gct gat ttg atg att 488
Thr Glu Ser Val Val Ser Gln Pro Pro Leu Thr Ala Asp Leu Met Ile
130                135                140                145

tgt cat agc gtt gaa gaa gct tca agg act caa atg gag gca aat ggg 536
Cys His Ser Val Glu Glu Ala Ser Arg Thr Gln Met Glu Ala Asn Gly
                150                155                160

ttg tgg aga aat gaa aca gga cag acc att gga ggg ttt gat ctg aat 584
Leu Trp Arg Asn Glu Thr Gly Gln Thr Ile Gly Gly Phe Asp Leu Asn
                165                170                175

tac gga att ggg ttt gat ttc aat ggt gtt cca gag att ggt ttt gga 632
Tyr Gly Ile Gly Phe Asp Phe Asn Gly Val Pro Glu Ile Gly Phe Gly
                180                185                190

gat aat caa acg cct gga ctt gaa tta agg ctg tct caa gtt ggg gtt 680
Asp Asn Gln Thr Pro Gly Leu Glu Leu Arg Leu Ser Gln Val Gly Val
                195                200                205

ttg aat cca cag gtt ttt caa caa atg ggt aaa gaa cag ttc agg gtt 728
Leu Asn Pro Gln Val Phe Gln Gln Met Gly Lys Glu Gln Phe Arg Val
210                215                . 220                225

ctt cat cat cat tca cat gaa gat cag cag cag agt gca gag gaa aat 776
Leu His His His Ser His Glu Asp Gln Gln Gln Ser Ala Glu Glu Asn
                230                235                240

ggt tca taa agctcaaacc tttaaattggc tttaaattag atatatttag 825
Gly Ser *

```

```

cgagaatgtg attgctagtg aaatctcaaa cttgatagct tattcatagc ttatcttact 885
tcgatcagca agagaacaat ggctcctaaa gctcaaacct tttttggcgt ttgctttaga 945
ttgtgcatac 955

```

```

<210> 182
<211> 243
<212> PRT
<213> Arabidopsis thaliana

```

```

<220>
<221> DOMAIN
<222> (64)...(129)
<223> Conserved domain

```

```

<400> 182
Met Val Met Glu Pro Lys Lys Asn Gln Asn Leu Pro Ser Phe Leu Asn
1          5          10          15
Pro Ser Arg Gln Asn Gln Asp Asn Asp Lys Lys Arg Lys Gln Thr Glu
20          25          30
Val Lys Gly Phe Asp Ile Val Val Gly Glu Lys Arg Lys Lys Lys Glu
35          40          45
Asn Glu Glu Glu Asp Gln Glu Ile Gln Ile Leu Tyr Glu Lys Glu Lys
50          55          60
Lys Lys Pro Asn Lys Asp Arg His Leu Lys Val Glu Gly Arg Gly Arg

```

```

65          70          75          80
Arg Val Arg Leu Pro Pro Leu Cys Ala Ala Arg Ile Tyr Gln Leu Thr
      85          90          95
Lys Glu Leu Gly His Lys Ser Asp Gly Glu Thr Leu Glu Trp Leu Leu
      100          105          110
Gln His Ala Glu Pro Ser Ile Leu Ser Ala Thr Val Asn Gly Ile Lys
      115          120          125
Pro Thr Glu Ser Val Val Ser Gln Pro Pro Leu Thr Ala Asp Leu Met
      130          135          140
Ile Cys His Ser Val Glu Glu Ala Ser Arg Thr Gln Met Glu Ala Asn
145          150          155          160
Gly Leu Trp Arg Asn Glu Thr Gly Gln Thr Ile Gly Gly Phe Asp Leu
      165          170          175
Asn Tyr Gly Ile Gly Phe Asp Phe Asn Gly Val Pro Glu Ile Gly Phe
      180          185          190
Gly Asp Asn Gln Thr Pro Gly Leu Glu Leu Arg Leu Ser Gln Val Gly
      195          200          205
Val Leu Asn Pro Gln Val Phe Gln Gln Met Gly Lys Glu Gln Phe Arg
210          215          220
Val Leu His His His Ser His Glu Asp Gln Gln Gln Ser Ala Glu Glu
225          230          235          240
Asn Gly Ser

```

<210> 183  
 <211> 883  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (42)...(764)

<400> 183  
 ctgaattcga actttggaag aagaagaagc ttgatcaat c atg gaa att gca acc 56  
 Met Glu Ile Ala Thr  
 1 5

```

gat aca gca aag cag atg aga gac gaa gag ttg ttc aaa gca gcg gaa 104
Asp Thr Ala Lys Gln Met Arg Asp Glu Glu Leu Phe Lys Ala Ala Glu
      10          15          20

```

```

tgg gga gat tca tcg ttg ttc atg tca tta tct gaa gaa cag ctc tct 152
Trp Gly Asp Ser Ser Leu Phe Met Ser Leu Ser Glu Glu Gln Leu Ser
      25          30          35

```

```

aaa tct ctc aat ttc aga aac gaa gat ggt cgc tct ctc ctc cat gtc 200
Lys Ser Leu Asn Phe Arg Asn Glu Asp Gly Arg Ser Leu Leu His Val
      40          45          50

```

```

gct gct tcc ttc ggc cat tct caa ata gtg aag ttg tta tca agt tca 248
Ala Ala Ser Phe Gly His Ser Gln Ile Val Lys Leu Leu Ser Ser Ser
      55          60          65

```

```

gat gaa gca aag act gta atc aat agc aag gat gat gaa gga tgg gct 296
Asp Glu Ala Lys Thr Val Ile Asn Ser Lys Asp Asp Glu Gly Trp Ala

```

```

70          75          80          85
cct ttg cat tcc gct gct agc atc ggt aat gct gag ctc gtt gag gtg 344
Pro Leu His Ser Ala Ala Ser Ile Gly Asn Ala Glu Leu Val Glu Val
          90          95          100

ctt ttg acc aga ggt gct gat gtc aat gcc aaa aat aac ggt ggt cgc 392
Leu Leu Thr Arg Gly Ala Asp Val Asn Ala Lys Asn Asn Gly Gly Arg
          105          110          115

act gct ctt cac tat gct gct agc aaa ggc cgg ttg gag att gct cag 440
Thr Ala Leu His Tyr Ala Ala Ser Lys Gly Arg Leu Glu Ile Ala Gln
          120          125          130

ctt tta tta aca cac ggt gca aag att aac atc aca gac aag gtt ggt 488
Leu Leu Leu Thr His Gly Ala Lys Ile Asn Ile Thr Asp Lys Val Gly
          135          140          145

tgc act ccg ctt cac agg gca gca agc gtg gga aag tta gaa gtt tgt 536
Cys Thr Pro Leu His Arg Ala Ala Ser Val Gly Lys Leu Glu Val Cys
          150          155          160          165

gaa ttt ctt att gaa gaa gga gca gag atc gat gct acg gat aaa atg 584
Glu Phe Leu Ile Glu Glu Gly Ala Glu Ile Asp Ala Thr Asp Lys Met
          170          175          180

ggt caa act gca ctc atg cat tca gtt atc tgc gat gac aaa cag gtt 632
Gly Gln Thr Ala Leu Met His Ser Val Ile Cys Asp Asp Lys Gln Val
          185          190          195

gcg ttc ctg ctt ata aga cat ggt gca gat gtg gat gta gaa gac aag 680
Ala Phe Leu Leu Ile Arg His Gly Ala Asp Val Asp Val Glu Asp Lys
          200          205          210

gaa ggc tac act gtt cta ggc cga gct acc aat gaa ttc cga cct gca 728
Glu Gly Tyr Thr Val Leu Gly Arg Ala Thr Asn Glu Phe Arg Pro Ala
          215          220          225

ctt atc gat gct gct aag gcc atg ctt gaa gga taa aatgactctg 774
Leu Ile Asp Ala Ala Lys Ala Met Leu Glu Gly *
          230          235          240

gattacttta aaacttacta actctgagag ttgtttagtt acttaaaagg atttttcttt 834
actgtatcat gtttgcaaaa tgtttctgcc ttatcaattc atgttctgt 883

<210> 184
<211> 240
<212> PRT
<213> Arabidopsis thaliana

<220>
<221> DOMAIN
<222> (65)...(228)
<223> Conserved domain

<400> 184
Met Glu Ile Ala Thr Asp Thr Ala Lys Gln Met Arg Asp Glu Glu Leu

```

```

1      5      10      15
Phe Lys Ala Ala Glu Trp Gly Asp Ser Ser Leu Phe Met Ser Leu Ser
20      25      30
Glu Glu Gln Leu Ser Lys Ser Leu Asn Phe Arg Asn Glu Asp Gly Arg
35      40      45
Ser Leu Leu His Val Ala Ala Ser Phe Gly His Ser Gln Ile Val Lys
50      55      60
Leu Leu Ser Ser Ser Asp Glu Ala Lys Thr Val Ile Asn Ser Lys Asp
65      70      75      80
Asp Glu Gly Trp Ala Pro Leu His Ser Ala Ala Ser Ile Gly Asn Ala
85      90      95
Glu Leu Val Glu Val Leu Leu Thr Arg Gly Ala Asp Val Asn Ala Lys
100      105      110
Asn Asn Gly Gly Arg Thr Ala Leu His Tyr Ala Ala Ser Lys Gly Arg
115      120      125
Leu Glu Ile Ala Gln Leu Leu Leu Thr His Gly Ala Lys Ile Asn Ile
130      135      140
Thr Asp Lys Val Gly Cys Thr Pro Leu His Arg Ala Ala Ser Val Gly
145      150      155      160
Lys Leu Glu Val Cys Glu Phe Leu Ile Glu Glu Gly Ala Glu Ile Asp
165      170      175
Ala Thr Asp Lys Met Gly Gln Thr Ala Leu Met His Ser Val Ile Cys
180      185      190
Asp Asp Lys Gln Val Ala Phe Leu Leu Ile Arg His Gly Ala Asp Val
195      200      205
Asp Val Glu Asp Lys Glu Gly Tyr Thr Val Leu Gly Arg Ala Thr Asn
210      215      220
Glu Phe Arg Pro Ala Leu Ile Asp Ala Ala Lys Ala Met Leu Glu Gly
225      230      235      240

```

&lt;210&gt; 185

&lt;211&gt; 1026

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1026)

&lt;400&gt; 185

```

atg gaa gga gga gga aga gga cca aat caa acg att ctc agt gaa ata 48
Met Glu Gly Gly Gly Arg Gly Pro Asn Gln Thr Ile Leu Ser Glu Ile
1      5      10      15

gaa cat atg cct gaa gct cca cgt caa cgt atc tct cat cac cgt cga 96
Glu His Met Pro Glu Ala Pro Arg Gln Arg Ile Ser His His Arg Arg
20      25      30

gct cgc tct gaa acc ttc ttc tcc ggc gaa tca atc gac gat ctc ctc 144
Ala Arg Ser Glu Thr Phe Phe Ser Gly Glu Ser Ile Asp Asp Leu Leu
35      40      45

tta ttc gat cct tcc gat atc gat ttc tct tct cta gac ttc ctc aac 192
Leu Phe Asp Pro Ser Asp Ile Asp Phe Ser Ser Leu Asp Phe Leu Asn
50      55      60

```

gct cca cca cca cca caa caa tca caa caa caa ccg caa gct tct ccc	240
Ala Pro Pro Pro Pro Gln Gln Ser Gln Gln Gln Pro Gln Ala Ser Pro	
65 70 75 80	
atg tcc gtt gat tcc gaa gaa acc tca tcc aac ggt gtt gtt cct cct	288
Met Ser Val Asp Ser Glu Glu Thr Ser Ser Asn Gly Val Val Pro Pro	
85 90 95	
aat tct ctt cct cca aaa ccc gaa gct aga ttc ggt cgc cat gtt cgt	336
Asn Ser Leu Pro Pro Lys Pro Glu Ala Arg Phe Gly Arg His Val Arg	
100 105 110	
agc ttc tcc gtt gat tcc gat ttc ttc gat gat ttg ggt gtt act gag	384
Ser Phe Ser Val Asp Ser Asp Phe Phe Asp Asp Leu Gly Val Thr Glu	
115 120 125	
gag aag ttt ata gct aca agt tca gga gag aag aag aaa ggg aat cat	432
Glu Lys Phe Ile Ala Thr Ser Ser Gly Glu Lys Lys Lys Gly Asn His	
130 135 140	
cat cat agc agg agt aat tct atg gat gga gag atg agt tcc gcg tcc	480
His His Ser Arg Ser Asn Ser Met Asp Gly Glu Met Ser Ser Ala Ser	
145 150 155 160	
ttt aat atc gaa tcc att tta gct tct gtg agt ggt aaa gat agt ggg	528
Phe Asn Ile Glu Ser Ile Leu Ala Ser Val Ser Gly Lys Asp Ser Gly	
165 170 175	
aag aag aat atg ggt atg ggt ggt gat aga ctt gct gag ctt gct ttg	576
Lys Lys Asn Met Gly Met Gly Gly Asp Arg Leu Ala Glu Leu Ala Leu	
180 185 190	
ctt gat cct aaa aga gct aaa agg att tta gcg aat aga caa tct gcg	624
Leu Asp Pro Lys Arg Ala Lys Arg Ile Leu Ala Asn Arg Gln Ser Ala	
195 200 205	
gcg agg tcc aaa gag agg aag att agg tat act ggt gag tta gag agg	672
Ala Arg Ser Lys Glu Arg Lys Ile Arg Tyr Thr Gly Glu Leu Glu Arg	
210 215 220	
aag gtt cag aca ctt cag aat gaa gct act aca ttg tct gct caa gtc	720
Lys Val Gln Thr Leu Gln Asn Glu Ala Thr Thr Leu Ser Ala Gln Val	
225 230 235 240	
act atg tta cag aga gga aca tca gag ctg aac act gaa aat aaa cac	768
Thr Met Leu Gln Arg Gly Thr Ser Glu Leu Asn Thr Glu Asn Lys His	
245 250 255	
ctc aaa atg cgg ctt caa gct tta gag caa caa gct gaa ctt agg gat	816
Leu Lys Met Arg Leu Gln Ala Leu Glu Gln Gln Ala Glu Leu Arg Asp	
260 265 270	
gct ttg aat gaa gcg ctg cgg gat gaa ctg aac cga ctt aag gtg gta	864
Ala Leu Asn Glu Ala Leu Arg Asp Glu Leu Asn Arg Leu Lys Val Val	
275 280 285	
gct gga gaa att cct cag ggg aat gga aat tct tac aac cgt gct caa	912

Ala Gly Glu Ile Pro Gln Gly Asn Gly Asn Ser Tyr Asn Arg Ala Gln  
 290 295 300

ttc tca tct cag caa tcg gca atg aat cag ttt ggg aac aaa acg aac 960  
 Phe Ser Ser Gln Gln Ser Ala Met Asn Gln Phe Gly Asn Lys Thr Asn  
 305 310 315 320

caa cag atg agt aca aac ggg cag cca tcg ctc cca agc tac atg gat 1008  
 Gln Gln Met Ser Thr Asn Gly Gln Pro Ser Leu Pro Ser Tyr Met Asp  
 325 330 335

ttc acc aag aga ggc tga 1026  
 Phe Thr Lys Arg Gly \*  
 340

<210> 186  
 <211> 341  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 186  
 Met Glu Gly Gly Gly Arg Gly Pro Asn Gln Thr Ile Leu Ser Glu Ile  
 1 5 10 15  
 Glu His Met Pro Glu Ala Pro Arg Gln Arg Ile Ser His His Arg Arg  
 20 25 30  
 Ala Arg Ser Glu Thr Phe Phe Ser Gly Glu Ser Ile Asp Asp Leu Leu  
 35 40 45  
 Leu Phe Asp Pro Ser Asp Ile Asp Phe Ser Ser Leu Asp Phe Leu Asn  
 50 55 60  
 Ala Pro Pro Pro Pro Gln Gln Ser Gln Gln Gln Pro Gln Ala Ser Pro  
 65 70 75 80  
 Met Ser Val Asp Ser Glu Glu Thr Ser Ser Asn Gly Val Val Pro Pro  
 85 90 95  
 Asn Ser Leu Pro Pro Lys Pro Glu Ala Arg Phe Gly Arg His Val Arg  
 100 105 110  
 Ser Phe Ser Val Asp Ser Asp Phe Phe Asp Asp Leu Gly Val Thr Glu  
 115 120 125  
 Glu Lys Phe Ile Ala Thr Ser Ser Gly Glu Lys Lys Lys Gly Asn His  
 130 135 140  
 His His Ser Arg Ser Asn Ser Met Asp Gly Glu Met Ser Ser Ala Ser  
 145 150 155 160  
 Phe Asn Ile Glu Ser Ile Leu Ala Ser Val Ser Gly Lys Asp Ser Gly  
 165 170 175  
 Lys Lys Asn Met Gly Met Gly Gly Asp Arg Leu Ala Glu Leu Ala Leu  
 180 185 190  
 Leu Asp Pro Lys Arg Ala Lys Arg Ile Leu Ala Asn Arg Gln Ser Ala  
 195 200 205  
 Ala Arg Ser Lys Glu Arg Lys Ile Arg Tyr Thr Gly Glu Leu Glu Arg  
 210 215 220  
 Lys Val Gln Thr Leu Gln Asn Glu Ala Thr Thr Leu Ser Ala Gln Val  
 225 230 235 240  
 Thr Met Leu Gln Arg Gly Thr Ser Glu Leu Asn Thr Glu Asn Lys His  
 245 250 255  
 Leu Lys Met Arg Leu Gln Ala Leu Glu Gln Gln Ala Glu Leu Arg Asp  
 260 265 270  
 Ala Leu Asn Glu Ala Leu Arg Asp Glu Leu Asn Arg Leu Lys Val Val



275	280	285
Ala Gly Glu Ile Pro Gln Gly Asn Gly Asn Ser Tyr Asn Arg Ala Gln		
290	295	300
Phe Ser Ser Gln Gln Ser Ala Met Asn Gln Phe Gly Asn Lys Thr Asn		
305	310	315
Gln Gln Met Ser Thr Asn Gly Gln Pro Ser Leu Pro Ser Tyr Met Asp		
	325	330
Phe Thr Lys Arg Gly		335
340		

&lt;210&gt; 187

&lt;211&gt; 1107

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (16)...(930)

&lt;400&gt; 187

aaaagatctg tttca atg gcg gat cgt gtt aaa ggt cca tgg agt caa gaa	51
Met Ala Asp Arg Val Lys Gly Pro Trp Ser Gln Glu	
1 5 10	
gaa gat gag cag cta cga agg atg gtt gag aaa tac gga ccg agg aat	99
Glu Asp Glu Gln Leu Arg Arg Met Val Glu Lys Tyr Gly Pro Arg Asn	
15 20 25	
tgg tct gcg att agc aaa tcg att cca ggt cga tct ggt aaa tcg tgt	147
Trp Ser Ala Ile Ser Lys Ser Ile Pro Gly Arg Ser Gly Lys Ser Cys	
30 35 40	
aga tta cgt tgg tgt aat cag tta tct ccg gag gtt gag cat cgt cct	195
Arg Leu Arg Trp Cys Asn Gln Leu Ser Pro Glu Val Glu His Arg Pro	
45 50 55 60	
ttc tcg ccg gag gaa gat gag act att gta acc gcc cgt gct cag ttt	243
Phe Ser Pro Glu Glu Asp Glu Thr Ile Val Thr Ala Arg Ala Gln Phe	
65 70 75	
ggt aac aag tgg gcg acg att gct cgt ctt ctt aac ggt cgt acg gat	291
Gly Asn Lys Trp Ala Thr Ile Ala Arg Leu Leu Asn Gly Arg Thr Asp	
80 85 90	
aac gcc gtt aaa aat cac tgg aac tct acg ctt aag agg aaa tgc agc	339
Asn Ala Val Lys Asn His Trp Asn Ser Thr Leu Lys Arg Lys Cys Ser	
95 100 105	
gga ggt gtg gcg gtt acg acg gtg acg gag acg gag gaa gat cag gat	387
Gly Gly Val Ala Val Thr Thr Val Thr Glu Thr Glu Glu Asp Gln Asp	
110 115 120	
cgg ccg aag aag agg aga tct gtt agc ttt gat cct gct ttt gct ccg	435
Arg Pro Lys Lys Arg Arg Ser Val Ser Phe Asp Pro Ala Phe Ala Pro	
125 130 135 140	

```

gtg gat act gga ttg tac atg agt cct gag agt cct aac gga atc gat 483
Val Asp Thr Gly Leu Tyr Met Ser Pro Glu Ser Pro Asn Gly Ile Asp
145 150 155

gtt agt gat tct agc acg att ccg tca ccg tcg tct cct gtt gct cag 531
Val Ser Asp Ser Ser Thr Ile Pro Ser Pro Ser Ser Pro Val Ala Gln
160 165 170

ctg ttt aaa cca atg ccg att tcc ggc ggt ttt acg gtg gtt ccg cag 579
Leu Phe Lys Pro Met Pro Ile Ser Gly Gly Phe Thr Val Val Pro Gln
175 180 185

ccg tta ccg gtt gaa atg tct tcg tct tcg gag gat cca cct act tcg 627
Pro Leu Pro Val Glu Met Ser Ser Ser Glu Asp Pro Pro Thr Ser
190 195 200

ttg agt ttg tca cta cct gga gct gag aac acg agt tcg agc cat aac 675
Leu Ser Leu Ser Leu Pro Gly Ala Glu Asn Thr Ser Ser Ser His Asn
205 210 215 220

aat aac aac aac gcg ttg atg ttt ccg aga ttt gag agt cag atg aag 723
Asn Asn Asn Asn Ala Leu Met Phe Pro Arg Phe Glu Ser Gln Met Lys
225 230 235

att aat gta gag gag aga gga gga gga gga gaa gga cgt aga ggt gag 771
Ile Asn Val Glu Glu Arg Gly Gly Gly Gly Glu Gly Arg Arg Gly Glu
240 245 250

ttt atg acg gtg gtg cag gag atg ata aaa gct gaa gtg agg agt tac 819
Phe Met Thr Val Val Gln Glu Met Ile Lys Ala Glu Val Arg Ser Tyr
255 260 265

atg gcg gaa atg cag aaa aca agt ggt gga ttc gtc gtc gga ggt tta 867
Met Ala Glu Met Gln Lys Thr Ser Gly Gly Phe Val Val Gly Gly Leu
270 275 280

tac gaa tcc ggc ggc aat ggt ggt ttt agg gat tgt gga gta ata aca 915
Tyr Glu Ser Gly Gly Asn Gly Gly Phe Arg Asp Cys Gly Val Ile Thr
285 290 295 300

cct aag gtt gag tag ttttggttta gggttaaaac ttgaatcgat tggggatttt 970
Pro Lys Val Glu *

```

```

caagagcatt catTTTTTggg gtttatggta aaattaaaaa caaaaacaaa atgtacagag 1030
gaattaaaat ttctatggaa taatcttaaa tctcaaatat ttgttacttg ttttggtgat 1090
tcataaccaa aatcaaa 1107

```

&lt;210&gt; 188

&lt;211&gt; 304

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (6)...(106)

&lt;223&gt; Conserved domain

&lt;400&gt; 188

```

Met Ala Asp Arg Val Lys Gly Pro Trp Ser Gln Glu Glu Asp Glu Gln
 1           5           10           15
Leu Arg Arg Met Val Glu Lys Tyr Gly Pro Arg Asn Trp Ser Ala Ile
           20           25           30
Ser Lys Ser Ile Pro Gly Arg Ser Gly Lys Ser Cys Arg Leu Arg Trp
           35           40           45
Cys Asn Gln Leu Ser Pro Glu Val Glu His Arg Pro Phe Ser Pro Glu
           50           55           60
Glu Asp Glu Thr Ile Val Thr Ala Arg Ala Gln Phe Gly Asn Lys Trp
65           70           75           80
Ala Thr Ile Ala Arg Leu Leu Asn Gly Arg Thr Asp Asn Ala Val Lys
           85           90           95
Asn His Trp Asn Ser Thr Leu Lys Arg Lys Cys Ser Gly Gly Val Ala
           100          105          110
Val Thr Thr Val Thr Glu Thr Glu Glu Asp Gln Asp Arg Pro Lys Lys
           115          120          125
Arg Arg Ser Val Ser Phe Asp Pro Ala Phe Ala Pro Val Asp Thr Gly
           130          135          140
Leu Tyr Met Ser Pro Glu Ser Pro Asn Gly Ile Asp Val Ser Asp Ser
145           150          155          160
Ser Thr Ile Pro Ser Pro Ser Ser Pro Val Ala Gln Leu Phe Lys Pro
           165          170          175
Met Pro Ile Ser Gly Gly Phe Thr Val Val Pro Gln Pro Leu Pro Val
           180          185          190
Glu Met Ser Ser Ser Glu Asp Pro Pro Thr Ser Leu Ser Leu Ser
           195          200          205
Leu Pro Gly Ala Glu Asn Thr Ser Ser Ser His Asn Asn Asn Asn
           210          215          220
Ala Leu Met Phe Pro Arg Phe Glu Ser Gln Met Lys Ile Asn Val Glu
225           230          235          240
Glu Arg Gly Gly Gly Gly Glu Gly Arg Arg Gly Glu Phe Met Thr Val
           245          250          255
Val Gln Glu Met Ile Lys Ala Glu Val Arg Ser Tyr Met Ala Glu Met
           260          265          270
Gln Lys Thr Ser Gly Gly Phe Val Val Gly Gly Leu Tyr Glu Ser Gly
           275          280          285
Gly Asn Gly Gly Phe Arg Asp Cys Gly Val Ile Thr Pro Lys Val Glu
290           295           300

```

&lt;210&gt; 189

&lt;211&gt; 1001

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (360)...(932)

&lt;400&gt; 189

```

tggcgacgcc ttgcgcgacc tctctcctc cctctgattc aaaccctaata tctgccgcta 60
ctccgcgcga tcagaaacaa ccacctctc ctctcaacc cacgaatcca tctctccac 120
caccgcacac caccgtcgta gctttagccg catcaacctc cgccgtcgcc agaaaaaccc 180
aacctgttct ctggactcaa gacgagaccc tctctttaat cgaatcatat aaagagaaat 240
ggttcgcaat tggctgtggt cctctaaaat caactcactg ggaagagatc gccgttgctg 300

```

```

cctcatctcg ttccggcggtt gaacggacat ctacacagtg taggcataag atcgagaag 359
atg cgt aaa aga ttc cga tct gag aga cag agt atg gga cct atc tca 407
Met Arg Lys Arg Phe Arg Ser Glu Arg Gln Ser Met Gly Pro Ile Ser
1 5 10 15

att tgg cct ttc tat aat caa atg gag gaa ttg gat tct tca aat cct 455
Ile Trp Pro Phe Tyr Asn Gln Met Glu Glu Leu Asp Ser Ser Asn Pro
20 25 30

gct ccg atc tca gct cgt cct ctc act cgt ctt cct cct aat tcg aat 503
Ala Pro Ile Ser Ala Arg Pro Leu Thr Arg Leu Pro Pro Asn Ser Asn
35 40 45

aat cgc tac gta gac gac gaa gaa gaa gat gaa gaa gag gat aac aat 551
Asn Arg Tyr Val Asp Asp Glu Glu Glu Asp Glu Glu Glu Asp Asn Asn
50 55 60

aac tac gaa gaa gag gaa gaa gaa gat gag aga cag agc aaa tcg agg 599
Asn Tyr Glu Glu Glu Glu Glu Glu Asp Glu Arg Gln Ser Lys Ser Arg
65 70 75 80

agt att aat tac ata ttg aga aga cct gga act gtt aac aga ttt gct 647
Ser Ile Asn Tyr Ile Leu Arg Arg Pro Gly Thr Val Asn Arg Phe Ala
85 90 95

gga gtt ggt gga gga ttg ttg tca tgg gga caa aaa gag aga tca tcg 695
Gly Val Gly Gly Gly Leu Leu Ser Trp Gly Gln Lys Glu Arg Ser Ser
100 105 110

aag agg aag agg aac gat gga gac ggt ggt gag agg cgg agg aaa gga 743
Lys Arg Lys Arg Asn Asp Gly Asp Gly Gly Glu Arg Arg Arg Lys Gly
115 120 125

atg cga gcc gtt gca gcg gag att agg gcg ttt gct gag aga gta atg 791
Met Arg Ala Val Ala Ala Glu Ile Arg Ala Phe Ala Glu Arg Val Met
130 135 140

gta atg gag aag aag aag att gag ttt gct aaa gag act gtg aga ctt 839
Val Met Glu Lys Lys Lys Ile Glu Phe Ala Lys Glu Thr Val Arg Leu
145 150 155 160

cgt aag gaa atg gag att cgt aga atc aat ttg att cag agt tct caa 887
Arg Lys Glu Met Glu Ile Arg Arg Ile Asn Leu Ile Gln Ser Ser Gln
165 170 175

act cag ctt ctt caa ttt atc aac aat gct ttt gat tct ttc taa 932
Thr Gln Leu Leu Gln Phe Ile Asn Asn Ala Phe Asp Ser Phe *
180 185 190

cgtagagatt tagtcattgt tgcttttaga tttgtaacat aggtaacgga tttgaagaaa 992
tgcaactatg 1001

<210> 190
<211> 190
<212> PRT
<213> Arabidopsis thaliana

```

<220>  
 <221> DOMAIN  
 <222> (107)...(181)  
 <223> Conserved domain

<400> 190  
 Met Arg Lys Arg Phe Arg Ser Glu Arg Gln Ser Met Gly Pro Ile Ser  
 1 5 10 15  
 Ile Trp Pro Phe Tyr Asn Gln Met Glu Glu Leu Asp Ser Ser Asn Pro  
 20 25 30  
 Ala Pro Ile Ser Ala Arg Pro Leu Thr Arg Leu Pro Pro Asn Ser Asn  
 35 40 45  
 Asn Arg Tyr Val Asp Asp Glu Glu Asp Glu Glu Asp Asn Asn  
 50 55 60  
 Asn Tyr Glu Glu Glu Glu Glu Glu Asp Glu Arg Gln Ser Lys Ser Arg  
 65 70 75 80  
 Ser Ile Asn Tyr Ile Leu Arg Arg Pro Gly Thr Val Asn Arg Phe Ala  
 85 90 95  
 Gly Val Gly Gly Gly Leu Leu Ser Trp Gly Gln Lys Glu Arg Ser Ser  
 100 105 110  
 Lys Arg Lys Arg Asn Asp Gly Asp Gly Gly Glu Arg Arg Arg Lys Gly  
 115 120 125  
 Met Arg Ala Val Ala Ala Glu Ile Arg Ala Phe Ala Glu Arg Val Met  
 130 135 140  
 Val Met Glu Lys Lys Lys Ile Glu Phe Ala Lys Glu Thr Val Arg Leu  
 145 150 155 160  
 Arg Lys Glu Met Glu Ile Arg Arg Ile Asn Leu Ile Gln Ser Ser Gln  
 165 170 175  
 Thr Gln Leu Leu Gln Phe Ile Asn Asn Ala Phe Asp Ser Phe  
 180 185 190

<210> 191  
 <211> 1317  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (19)...(1314)

<400> 191  
 tcgacagttg gtgttttaa atg gaa ttc tca ttg atc tcc gat gaa aat ttc 51  
 Met Glu Phe Ser Leu Ile Ser Asp Glu Asn Phe  
 1 5 10  
 atc tta aac gat gct cat cga tgt cct gtt aat gat gat gca tcg cat 99  
 Ile Leu Asn Asp Ala His Arg Cys Pro Val Asn Asp Asp Ala Ser His  
 15 20 25  
 ata cca aaa aat aat ctt aac ttt ttc aat gac aat tta ggt caa tct 147  
 Ile Pro Lys Asn Asn Leu Asn Phe Phe Asn Asp Asn Leu Gly Gln Ser  
 30 35 40  
 tct cgt act gga tgg tct ttt tca cca gat ctg act gat atc tct aat 195  
 Ser Arg Thr Gly Trp Ser Phe Ser Pro Asp Leu Thr Asp Ile Ser Asn  
 45 50 55

cag cac cac caa aat ctt ata cct ttg atc cca aat tat gat tct caa	243
Gln His His Gln Asn Leu Ile Pro Leu Ile Pro Asn Tyr Asp Ser Gln	
60 65 70 75	
aat cag aat ttg gat act aat caa aac cat ttg gtt tac aac tct tcc	291
Asn Gln Asn Leu Asp Thr Asn Gln Asn His Leu Val Tyr Asn Ser Ser	
80 85 90	
tcc tat gaa att cca tca aac tac cca ttc atg agt atc aaa tcc tac	339
Ser Tyr Glu Ile Pro Ser Asn Tyr Pro Phe Met Ser Ile Lys Ser Tyr	
95 100 105	
tca aat att gat act ctt gag caa tct atg aat aac ata gtc aac aac	387
Ser Asn Ile Asp Thr Leu Glu Gln Ser Met Asn Asn Ile Val Asn Asn	
110 115 120	
gga aaa atc cat atg atc gat aat ccc cca att ttt gca aac cct aag	435
Gly Lys Ile His Met Ile Asp Asn Pro Pro Ile Phe Ala Asn Pro Lys	
125 130 135	
ggg ata ttc gaa aat ttt cat gat ttg cag gaa tac act ata ggg aat	483
Gly Ile Phe Glu Asn Phe His Asp Leu Gln Glu Tyr Thr Ile Gly Asn	
140 145 150 155	
gag att gta cac aat gaa gaa ctg act aat aag ggt tat gag cca act	531
Glu Ile Val His Asn Glu Glu Leu Thr Asn Lys Gly Tyr Glu Pro Thr	
160 165 170	
ctt gat aag gta atg ggt gaa cct caa tta ttt gat gtg cca gtg cta	579
Leu Asp Lys Val Met Gly Glu Pro Gln Leu Phe Asp Val Pro Val Leu	
175 180 185	
gaa ggt atc aaa aat aca acc aac gaa atc atg aat caa ctt gaa gat	627
Glu Gly Ile Lys Asn Thr Thr Asn Glu Ile Met Asn Gln Leu Glu Asp	
190 195 200	
gat aag atg aaa aaa act tat gag aat aaa aag gag gcg agc aca agc	675
Asp Lys Met Lys Lys Thr Tyr Glu Asn Lys Lys Glu Ala Ser Thr Ser	
205 210 215	
aaa tat ttg aag aaa agt gac atc acc aaa aaa aga tgg aca gaa tct	723
Lys Tyr Leu Lys Lys Ser Asp Ile Thr Lys Lys Arg Trp Thr Glu Ser	
220 225 230 235	
gag gac ata aaa ttg aaa gag atg gtg gca ctt gaa ccc aaa aaa tgg	771
Glu Asp Ile Lys Leu Lys Glu Met Val Ala Leu Glu Pro Lys Lys Trp	
240 245 250	
aca aag gtc gca aaa cat ttt gaa gga cga aca cca aaa caa tgc aga	819
Thr Lys Val Ala Lys His Phe Glu Gly Arg Thr Pro Lys Gln Cys Arg	
255 260 265	
gag agg tgg cat aac cat gct cgt cct aac gtt aag aaa act act tgg	867
Glu Arg Trp His Asn His Ala Arg Pro Asn Val Lys Lys Thr Thr Trp	
270 275 280	

agt gaa gaa gaa gat caa ata cta att gaa gta cac aaa gta att ggt 915  
 Ser Glu Glu Glu Asp Gln Ile Leu Ile Glu Val His Lys Val Ile Gly  
 285 290 295  
 gcc aaa tgg atc cag atc tca gag caa ctt ccg ggg agg agc tat aac 963  
 Ala Lys Trp Ile Gln Ile Ser Glu Gln Leu Pro Gly Arg Ser Tyr Asn  
 300 305 310 315  
 aat gtt aag aac cat tgg aat acc aca aag cgt cgg gtg caa aat aaa 1011  
 Asn Val Lys Asn His Trp Asn Thr Thr Lys Arg Arg Val Gln Asn Lys  
 320 325 330  
 agt ggt cga act gtc aac cgt gtc ggc aat aac atc ctt gaa aac tac 1059  
 Ser Gly Arg Thr Val Asn Arg Val Gly Asn Asn Ile Leu Glu Asn Tyr  
 335 340 345  
 ata agg agt atc aca atc aac aac gat gat gaa tct gat ggt gaa ccg 1107  
 Ile Arg Ser Ile Thr Ile Asn Asn Asp Asp Glu Ser Asp Gly Glu Pro  
 350 355 360  
 aca aac att gag aat tat cat gat gac tct gaa gat atg tta tat ggg 1155  
 Thr Asn Ile Glu Asn Tyr His Asp Asp Ser Glu Asp Met Leu Tyr Gly  
 365 370 375  
 gag atg aat ttg agt cca gag gct ata acg cag acc acc aaa cca ttg 1203  
 Glu Met Asn Leu Ser Pro Glu Ala Ile Thr Gln Thr Thr Lys Pro Leu  
 380 385 390 395  
 acc gat gct tct aca ata tca cct tat att ccg atg cca aaa gaa aac 1251  
 Thr Asp Ala Ser Thr Ile Ser Pro Tyr Ile Pro Met Pro Lys Glu Asn  
 400 405 410  
 tat act ctg gag gtt tgt gaa tca ctg gag gac tat cta gag ctg ctg 1299  
 Tyr Thr Leu Glu Val Cys Glu Ser Leu Glu Asp Tyr Leu Glu Leu Leu  
 415 420 425  
 cgt tgg tgg gat taa tgc 1317  
 Arg Trp Trp Asp \*  
 430

<210> 192  
 <211> 431  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (264)...(332)  
 <223> Conserved domain

<400> 192  
 Met Glu Phe Ser Leu Ile Ser Asp Glu Asn Phe Ile Leu Asn Asp Ala  
 1 5 10 15  
 His Arg Cys Pro Val Asn Asp Asp Ala Ser His Ile Pro Lys Asn Asn  
 20 25 30  
 Leu Asn Phe Phe Asn Asp Asn Leu Gly Gln Ser Ser Arg Thr Gly Trp

		35					40					45				
Ser	Phe	Ser	Pro	Asp	Leu	Thr	Asp	Ile	Ser	Asn	Gln	His	His	Gln	Asn	
	50					55					60					
Leu	Ile	Pro	Leu	Ile	Pro	Asn	Tyr	Asp	Ser	Gln	Asn	Gln	Asn	Leu	Asp	
65					70					75					80	
Thr	Asn	Gln	Asn	His	Leu	Val	Tyr	Asn	Ser	Ser	Ser	Tyr	Glu	Ile	Pro	
				85					90					95		
Ser	Asn	Tyr	Pro	Phe	Met	Ser	Ile	Lys	Ser	Tyr	Ser	Asn	Ile	Asp	Thr	
			100					105					110			
Leu	Glu	Gln	Ser	Met	Asn	Asn	Ile	Val	Asn	Asn	Gly	Lys	Ile	His	Met	
		115					120					125				
Ile	Asp	Asn	Pro	Pro	Ile	Phe	Ala	Asn	Pro	Lys	Gly	Ile	Phe	Glu	Asn	
	130					135					140					
Phe	His	Asp	Leu	Gln	Glu	Tyr	Thr	Ile	Gly	Asn	Glu	Ile	Val	His	Asn	
145				150					155						160	
Glu	Glu	Leu	Thr	Asn	Lys	Gly	Tyr	Glu	Pro	Thr	Leu	Asp	Lys	Val	Met	
				165					170					175		
Gly	Glu	Pro	Gln	Leu	Phe	Asp	Val	Pro	Val	Leu	Glu	Gly	Ile	Lys	Asn	
			180				185						190			
Thr	Thr	Asn	Glu	Ile	Met	Asn	Gln	Leu	Glu	Asp	Asp	Lys	Met	Lys	Lys	
		195					200					205				
Thr	Tyr	Glu	Asn	Lys	Lys	Glu	Ala	Ser	Thr	Ser	Lys	Tyr	Leu	Lys	Lys	
	210					215					220					
Ser	Asp	Ile	Thr	Lys	Lys	Arg	Trp	Thr	Glu	Ser	Glu	Asp	Ile	Lys	Leu	
225				230					235						240	
Lys	Glu	Met	Val	Ala	Leu	Glu	Pro	Lys	Lys	Trp	Thr	Lys	Val	Ala	Lys	
				245					250					255		
His	Phe	Glu	Gly	Arg	Thr	Pro	Lys	Gln	Cys	Arg	Glu	Arg	Trp	His	Asn	
			260					265					270			
His	Ala	Arg	Pro	Asn	Val	Lys	Lys	Thr	Thr	Trp	Ser	Glu	Glu	Glu	Asp	
	275						280					285				
Gln	Ile	Leu	Ile	Glu	Val	His	Lys	Val	Ile	Gly	Ala	Lys	Trp	Ile	Gln	
	290					295					300					
Ile	Ser	Glu	Gln	Leu	Pro	Gly	Arg	Ser	Tyr	Asn	Asn	Val	Lys	Asn	His	
305				310					315						320	
Trp	Asn	Thr	Thr	Lys	Arg	Arg	Val	Gln	Asn	Lys	Ser	Gly	Arg	Thr	Val	
				325					330					335		
Asn	Arg	Val	Gly	Asn	Asn	Ile	Leu	Glu	Asn	Tyr	Ile	Arg	Ser	Ile	Thr	
			340					345					350			
Ile	Asn	Asn	Asp	Asp	Glu	Ser	Asp	Gly	Glu	Pro	Thr	Asn	Ile	Glu	Asn	
		355					360					365				
Tyr	His	Asp	Asp	Ser	Glu	Asp	Met	Leu	Tyr	Gly	Glu	Met	Asn	Leu	Ser	

<210> 193

<211> 1167

<212> DNA

<213> Arabidopsis thaliana

 $\langle 220 \rangle$



&lt;221&gt; CDS

&lt;222&gt; (177)...(956)

&lt;400&gt; 193

ctgtttttgt atccgtgtaa attaatcaca cggtagtttt tgatgaaaag acaacaatcg 60  
 gagaacaatc tgggtctgctg ctaaaattta ataaattggt ttgtctaatt gtctccaccc 120  
 ataaaaaagc gcgaattcaa ttcaccgact aaagacattc tccggtggag accccg atg 179  
 Met

1

caa tcc act cat ata agc ggc gga agt agc ggt ggt ggt ggt gga gga 227  
 Gln Ser Thr His Ile Ser Gly Gly Ser Ser Gly Gly Gly Gly Gly Gly  
 5 10 15

gga gga gag gtg agt cga agt gga tta tct cgg atc cgt tca gct cca 275  
 Gly Gly Glu Val Ser Arg Ser Gly Leu Ser Arg Ile Arg Ser Ala Pro  
 20 25 30

gct act tgg att gaa acc cta ctc gaa gaa gat gaa gaa gaa ggt tta 323  
 Ala Thr Trp Ile Glu Thr Leu Leu Glu Glu Asp Glu Glu Glu Gly Leu  
 35 40 45

aaa cct aac ctt tgt tta aca gag ctg ctt act ggt aat aat aac tct 371  
 Lys Pro Asn Leu Cys Leu Thr Glu Leu Leu Thr Gly Asn Asn Asn Ser  
 50 55 60 65

gga gga gtg ata acg agt cgt gac gac tcg ttc gag ttc ctg agt tct 419  
 Gly Gly Val Ile Thr Ser Arg Asp Asp Phe Glu Phe Leu Ser Ser  
 70 75 80

gtt gag caa gga ttg tat aat cat cat caa ggt ggt ggc ttt cac cgt 467  
 Val Glu Gln Gly Leu Tyr Asn His His Gln Gly Gly Gly Phe His Arg  
 85 90 95

cag aat agt tct ccg gct gat ttt ctt agt ggg tct ggt tct ggg act 515  
 Gln Asn Ser Ser Pro Ala Asp Phe Leu Ser Gly Ser Gly Ser Gly Thr  
 100 105 110

gat ggg tat ttc tct aat ttt ggt att ccg gcg aat tat gac tat ttg 563  
 Asp Gly Tyr Phe Ser Asn Phe Gly Ile Pro Ala Asn Tyr Asp Tyr Leu  
 115 120 125

tcg acc aac gtt gat att tct ccg act aaa cgg tct aga gat atg gaa 611  
 Ser Thr Asn Val Asp Ile Ser Pro Thr Lys Arg Ser Arg Asp Met Glu  
 130 135 140 145

aca cag ttt tct tct cag ctg aaa gaa gag caa atg agt ggt ggg ata 659  
 Thr Gln Phe Ser Ser Gln Leu Lys Glu Glu Gln Met Ser Gly Gly Ile  
 150 155 160

tca gga atg atg gat atg aac atg gac aag att ttt gag gat tca gtt 707  
 Ser Gly Met Met Asp Met Asn Met Asp Lys Ile Phe Glu Asp Ser Val  
 165 170 175

cct tgt agg gtt cgt gct aaa cgt ggt tgt gct act cat cct cgt agc 755  
 Pro Cys Arg Val Arg Ala Lys Arg Gly Cys Ala Thr His Pro Arg Ser  
 180 185 190

```

att gct gaa cgg gtg aga aga acg cga ata agt gat cgg att agg agg 803
Ile Ala Glu Arg Val Arg Arg Thr Arg Ile Ser Asp Arg Ile Arg Arg
195 200 205

ctg caa gag ctt gtt cct aac atg gat aag caa acc aac act gca gac 851
Leu Gln Glu Leu Val Pro Asn Met Asp Lys Gln Thr Asn Thr Ala Asp
210 215 220 225

atg ttg gaa gaa gct gtg gag tat gtg aag gct ctt caa agc cag atc 899
Met Leu Glu Glu Ala Val Glu Tyr Val Lys Ala Leu Gln Ser Gln Ile
230 235 240

cag gaa ttg aca gag cag cag aag aga tgc aaa tgc aaa cct aaa gaa 947
Gln Glu Leu Thr Glu Gln Gln Lys Arg Cys Lys Cys Lys Pro Lys Glu
245 250 255

gaa caa taa tgtatccttt aggatttgat atatctgtat tttatttttg 996
Glu Gln *

```

```

tactatctaa aaatgggtgat gatctgttcg aaaattcgaa acatgatctt atatattgaa 1056
ctagaaaaaaa tagatatata tgaatttttag ctgtaaaatt tttgtacaat aaggagaaaa 1116
agatttagaa gagtcaataa aaagatgatg tttacaagtc aaaaaaaaaa a 1167

```

```

<210> 194
<211> 259
<212> PRT
<213> Arabidopsis thaliana

```

```

<220>
<221> DOMAIN
<222> (175)...(245)
<223> Conserved domain

```

```

<400> 194
Met Gln Ser Thr His Ile Ser Gly Gly Ser Ser Gly Gly Gly Gly Gly
1 5 10 15
Gly Gly Gly Glu Val Ser Arg Ser Gly Leu Ser Arg Ile Arg Ser Ala
20 25 30
Pro Ala Thr Trp Ile Glu Thr Leu Leu Glu Glu Asp Glu Glu Gly
35 40 45
Leu Lys Pro Asn Leu Cys Leu Thr Glu Leu Leu Thr Gly Asn Asn Asn
50 55 60
Ser Gly Gly Val Ile Thr Ser Arg Asp Asp Ser Phe Glu Phe Leu Ser
65 70 75 80
Ser Val Glu Gln Gly Leu Tyr Asn His His Gln Gly Gly Gly Phe His
85 90 95
Arg Gln Asn Ser Ser Pro Ala Asp Phe Leu Ser Gly Ser Gly Ser Gly
100 105 110
Thr Asp Gly Tyr Phe Ser Asn Phe Gly Ile Pro Ala Asn Tyr Asp Tyr
115 120 125
Leu Ser Thr Asn Val Asp Ile Ser Pro Thr Lys Arg Ser Arg Asp Met
130 135 140
Glu Thr Gln Phe Ser Ser Gln Leu Lys Glu Glu Gln Met Ser Gly Gly
145 150 155 160
Ile Ser Gly Met Met Asp Met Asn Met Asp Lys Ile Phe Glu Asp Ser

```

				165				170				175			
Val	Pro	Cys	Arg	Val	Arg	Ala	Lys	Arg	Gly	Cys	Ala	Thr	His	Pro	Arg
180				185				190							
Ser	Ile	Ala	Glu	Arg	Val	Arg	Arg	Thr	Arg	Ile	Ser	Asp	Arg	Ile	Arg
195				200				205							
Arg	Leu	Gln	Glu	Leu	Val	Pro	Asn	Met	Asp	Lys	Gln	Thr	Asn	Thr	Ala
210				215				220							
Asp	Met	Leu	Glu	Glu	Ala	Val	Glu	Tyr	Val	Lys	Ala	Leu	Gln	Ser	Gln
225				230				235				240			
Ile	Gln	Glu	Leu	Thr	Glu	Gln	Gln	Lys	Arg	Cys	Lys	Cys	Lys	Pro	Lys
245				250				255							
Glu	Glu	Gln													

```
<210> 195
<211> 964
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>  
<221> CDS  
<222> (63) ... (869)
```

[illegible]

```

gtc aag aag gag aag acg agt cct gtt tcg gcg gcg gtg acg gcg gcg 491
Val Lys Lys Glu Lys Thr Ser Pro Val Ser Ala Ala Val Thr Ala Ala
      130                      135                      140

aag gga aag cat tat aga gga gtg aga caa agg ccg tgg ggg aaa ttt 539
Lys Gly Lys His Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Phe
      145                      150                      155

gcg gcg gag att aga gat ccg gcg aag aac gga gct agg gtt tgg tta 587
Ala Ala Glu Ile Arg Asp Pro Ala Lys Asn Gly Ala Arg Val Trp Leu
      160                      165                      170                      175

gga acg ttt gag acg gcg gag gac gcg gcg ttg gct tac gac aga gct 635
Gly Thr Phe Glu Thr Ala Glu Asp Ala Leu Ala Tyr Asp Arg Ala
      180                      185                      190

gct ttc agg atg cgt ggt tcc cgc gct ttg ttg aat ttt ccg ttg aga 683
Ala Phe Arg Met Arg Gly Ser Arg Ala Leu Leu Asn Phe Pro Leu Arg
      195                      200                      205

gtt aat tca gga gaa ccc gac ccg gtt cga atc aag tcc aag aga tct 731
Val Asn Ser Gly Glu Pro Asp Pro Val Arg Ile Lys Ser Lys Arg Ser
      210                      215                      220

tct ttt tct tct tct aac gag aac gga gct ccg aag aag agg aga acg 779
Ser Phe Ser Ser Ser Asn Glu Asn Gly Ala Pro Lys Lys Arg Arg Thr
      225                      230                      235

gtg gcc gcc ggt ggt gga atg gat aag gga ttg acg gtg aag tgc gag 827
Val Ala Ala Gly Gly Gly Met Asp Lys Gly Leu Thr Val Lys Cys Glu
      240                      245                      250                      255

gtt gtt gaa gtg gca cgt ggc gat cgt tta ttg gtt tta taa 869
Val Val Glu Val Ala Arg Gly Asp Arg Leu Leu Val Leu *
      260                      265

ttttgatttt tctttgttgg atgattatat gattcttcaa aaaagaagaa cgtaataaaa 929
aaaattcgtt tattattaaa aaaaaaaaaa aaaaaa 964

```

&lt;210&gt; 196

&lt;211&gt; 268

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (145)...(213)

&lt;223&gt; Conserved domain

&lt;400&gt; 196

```

Met Ser Met Thr Ala Asp Ser Gln Ser Asp Tyr Ala Phe Leu Glu Ser
1      5      10      15
Ile Arg Arg His Leu Leu Gly Glu Ser Glu Pro Ile Leu Ser Glu Ser
      20      25      30
Thr Ala Ser Ser Val Thr Gln Ser Cys Val Thr Gly Gln Ser Ile Lys
      35      40      45

```

Pro Val Tyr Gly Arg Asn Pro Ser Phe Ser Lys Leu Tyr Pro Cys Phe  
 50 55 60  
 Thr Glu Ser Trp Gly Asp Leu Pro Leu Lys Glu Asn Asp Ser Glu Asp  
 65 70 75 80  
 Met Leu Val Tyr Gly Ile Leu Asn Asp Ala Phe His Gly Gly Trp Glu  
 85 90 95  
 Pro Ser Ser Ser Ser Asp Glu Asp Arg Ser Ser Phe Pro Ser Val  
 100 105 110  
 Lys Ile Glu Thr Pro Glu Ser Phe Ala Ala Val Asp Ser Val Pro Val  
 115 120 125  
 Lys Lys Glu Lys Thr Ser Pro Val Ser Ala Ala Val Thr Ala Ala Lys  
 130 135 140  
 Gly Lys His Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Phe Ala  
 145 150 155 160  
 Ala Glu Ile Arg Asp Pro Ala Lys Asn Gly Ala Arg Val Trp Leu Gly  
 165 170 175  
 Thr Phe Glu Thr Ala Glu Asp Ala Ala Leu Ala Tyr Asp Arg Ala Ala  
 180 185 190  
 Phe Arg Met Arg Gly Ser Arg Ala Leu Leu Asn Phe Pro Leu Arg Val  
 195 200 205  
 Asn Ser Gly Glu Pro Asp Pro Val Arg Ile Lys Ser Lys Arg Ser Ser  
 210 215 220  
 Phe Ser Ser Ser Asn Glu Asn Gly Ala Pro Lys Lys Arg Arg Thr Val  
 225 230 235 240  
 Ala Ala Gly Gly Gly Met Asp Lys Gly Leu Thr Val Lys Cys Glu Val  
 245 250 255  
 Val Glu Val Ala Arg Gly Asp Arg Leu Leu Val Leu  
 260 265

&lt;210&gt; 197

&lt;211&gt; 582

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(582)

&lt;400&gt; 197

atg gag att gag aag gat gag gac gac aca aca ttg gtt gat tct gga 48  
 Met Glu Ile Glu Lys Asp Glu Asp Asp Thr Thr Leu Val Asp Ser Gly  
 1 5 10 15  
 gga gac ttc gac tgc aac ata tgt ttg gat cag gtt cga gac ccg gtc 96  
 Gly Asp Phe Asp Cys Asn Ile Cys Leu Asp Gln Val Arg Asp Pro Val  
 20 25 30  
 gtg act tta tgt ggc cac ctg ttt tgt tgg ccc tgc att cac aag tgg 144  
 Val Thr Leu Cys Gly His Leu Phe Cys Trp Pro Cys Ile His Lys Trp  
 35 40 45  
 act tat gcg tcc aac aat tca aga caa cga gtc gat caa tac gat cat 192  
 Thr Tyr Ala Ser Asn Asn Ser Arg Gln Arg Val Asp Gln Tyr Asp His  
 50 55 60  
 aag agg gaa cca cca aaa tgt ccg gta tgc aaa tct gat gtc tcc gag 240

```

Lys Arg Glu Pro Pro Lys Cys Pro Val Cys Lys Ser Asp Val Ser Glu
65              70              75              80

gct acg ctt gtc ccg atc tac gga cga gga cag aaa gct ccc cag tcc 288
Ala Thr Leu Val Pro Ile Tyr Gly Arg Gly Gln Lys Ala Pro Gln Ser
              85              90              95

ggt tca aat gta ccg agc aga cca act ggt ccg gtt tat gac tta aga 336
Gly Ser Asn Val Pro Ser Arg Pro Thr Gly Pro Val Tyr Asp Leu Arg
              100              105              110

gga gtt ggt caa cgt tta gga gaa ggg gag agt caa cgt tac atg tat 384
Gly Val Gly Gln Arg Leu Gly Glu Gly Glu Ser Gln Arg Tyr Met Tyr
              115              120              125

aga atg cct gat ccg gtg atg ggt gtg gta tgc gaa atg gta tac ccg 432
Arg Met Pro Asp Pro Val Met Gly Val Val Cys Glu Met Val Tyr Arg
              130              135              140

aga cta ttt gga gag tct tcg agc aac atg gca cct tac cgc gat atg 480
Arg Leu Phe Gly Glu Ser Ser Asn Met Ala Pro Tyr Arg Asp Met
145              150              155              160

aat gtc cgg tct agg cga cgg gca atg cag gct gag gag tca tta agc 528
Asn Val Arg Ser Arg Arg Arg Ala Met Gln Ala Glu Glu Ser Leu Ser
              165              170              175

aga gtc tac ttg ttt cta ctt tgc ttc atg ttt atg tgt cta ttt ctc 576
Arg Val Tyr Leu Phe Leu Leu Cys Phe Met Phe Met Cys Leu Phe Leu
              180              185              190

ttc taa
Phe * 582

```

```

<210> 198
<211> 193
<212> PRT
<213> Arabidopsis thaliana

```

```

<220>
<221> DOMAIN
<222> (21)...(74).
<223> Conserved domain

```

```

<400> 198
Met Glu Ile Glu Lys Asp Glu Asp Asp Thr Thr Leu Val Asp Ser Gly
1      5      10
Gly Asp Phe Asp Cys Asn Ile Cys Leu Asp Gln Val Arg Asp Pro Val
20     25     30
Val Thr Leu Cys Gly His Leu Phe Cys Trp Pro Cys Ile His Lys Trp
35     40     45
Thr Tyr Ala Ser Asn Asn Ser Arg Gln Arg Val Asp Gln Tyr Asp His
50     55     60
Lys Arg Glu Pro Pro Lys Cys Pro Val Cys Lys Ser Asp Val Ser Glu
65     70     75     80

```

Ala Thr Leu Val Pro Ile Tyr Gly Arg Gly Gln Lys Ala Pro Gln Ser  
                             85                            90                            95  
 Gly Ser Asn Val Pro Ser Arg Pro Thr Gly Pro Val Tyr Asp Leu Arg  
                             100                            105                            110  
 Gly Val Gly Gln Arg Leu Gly Glu Gly Glu Ser Gln Arg Tyr Met Tyr  
                             115                            120                            125  
 Arg Met Pro Asp Pro Val Met Gly Val Val Cys Glu Met Val Tyr Arg  
                             130                            135                            140  
 Arg Leu Phe Gly Glu Ser Ser Ser Asn Met Ala Pro Tyr Arg Asp Met  
                             145                            150                            155                            160  
 Asn Val Arg Ser Arg Arg Arg Ala Met Gln Ala Glu Glu Ser Leu Ser  
                             165                            170                            175  
 Arg Val Tyr Leu Phe Leu Leu Cys Phe Met Phe Met Cys Leu Phe Leu  
                             180                            185                            190  
 Phe

&lt;210&gt; 199

&lt;211&gt; 952

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (65)...(883)

&lt;400&gt; 199

tcgaaatcat caagaatctt gtctttaatt tccaattggg tttttcaaaa cccgttacct 60  
 cagg atg atg aac ccg ttt ctc ccg gaa ggc tgc gat cca cca cca cca 109  
     Met Met Asn Pro Phe Leu Pro Glu Gly Cys Asp Pro Pro Pro Pro  
       1                            5                            10                            15

cca caa cca atg gag ggt tta cac gaa aat gct cca cct cca ttt ctg 157  
 Pro Gln Pro Met Glu Gly Leu His Glu Asn Ala Pro Pro Pro Phe Leu  
                             20                            25                            30

acc aag aca ttt gag atg gtg gat gat cca aac act gac cac atc gta 205  
 Thr Lys Thr Phe Glu Met Val Asp Asp Pro Asn Thr Asp His Ile Val  
                             35                            40                            45

tct tgg aac aga gga gga aca agt ttt gtc gtc tgg gat ttg cat tct 253  
 Ser Trp Asn Arg Gly Gly Thr Ser Phe Val Val Trp Asp Leu His Ser  
                             50                            55                            60

ttc tcc acg att ctc ctt cct cgt cat ttc aaa cac agc aat ttc tca 301  
 Phe Ser Thr Ile Leu Leu Pro Arg His Phe Lys His Ser Asn Phe Ser  
                             65                            70                            75

agt ttc atc aga caa ctc aat act tat ggt ttc aga aag ata gaa gca 349  
 Ser Phe Ile Arg Gln Leu Asn Thr Tyr Gly Phe Arg Lys Ile Glu Ala  
                             80                            85                            90                            95

gag aga tgg gaa ttt gca aac gaa gag ttt ttg tta gga caa aga cag 397  
 Glu Arg Trp Glu Phe Ala Asn Glu Glu Phe Leu Leu Gly Gln Arg Gln  
                             100                            105                            110

```

ttg ttg aag aac atc aag agg aga aac cct ttt act cca tca tct tca 445
Leu Leu Lys Asn Ile Lys Arg Arg Asn Pro Phe Thr Pro Ser Ser Ser
115 120 125

cca agc cat gac gct tgc aac gag ctt cgc aga gag aag caa gtg cta 493
Pro Ser His Asp Ala Cys Asn Glu Leu Arg Arg Glu Lys Gln Val Leu
130 135 140

atg atg gag ata gtg agt cta aga cag cag caa caa aca acg aaa agc 541
Met Met Glu Ile Val Ser Leu Arg Gln Gln Gln Gln Thr Thr Lys Ser
145 150 155

tac atc aaa gct atg gaa cag agg ata gaa gga aca gag agg aaa cag 589
Tyr Ile Lys Ala Met Glu Gln Arg Ile Glu Gly Thr Glu Arg Lys Gln
160 165 170 175

aga caa atg atg tcg ttt ctg gct aga gca atg cag agt cct tcg ttt 637
Arg Gln Met Met Ser Phe Leu Ala Arg Ala Met Gln Ser Pro Ser Phe
180 185 190

ttg cat cag ttg ttg aaa caa aga gat aaa aaa att aag gag ctt gag 685
Leu His Gln Leu Leu Lys Gln Arg Asp Lys Lys Ile Lys Glu Leu Glu
195 200 205

gat aat gag tca gca aag agg aaa aga ggt tct tct tcg atg tcg gaa 733
Asp Asn Glu Ser Ala Lys Arg Lys Arg Gly Ser Ser Ser Met Ser Glu
210 215 220

ttg gaa gtt ttg gct ttg gag atg caa ggg cat gga aaa cag agg aat 781
Leu Glu Val Leu Ala Leu Glu Met Gln Gly His Gly Lys Gln Arg Asn
225 230 235

atg ttg gaa gaa gag gat cat caa ctg gtg gta gag aga gag ttg gat 829
Met Leu Glu Glu Glu Asp His Gln Leu Val Val Glu Arg Glu Leu Asp
240 245 250 255

gat ggt ttc tgg gaa gag ttg ctt agt gat gag agt ttg gct tcc acc 877
Asp Gly Phe Trp Glu Glu Leu Leu Ser Asp Glu Ser Leu Ala Ser Thr
260 265 270

tcc taa ctagatggat ttcttttttg ttttggtttt agttttcttct actttcaagc 933
Ser *

tcattttctt ctgtcacaa 952

```

&lt;210&gt; 200

&lt;211&gt; 272

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (31)...(120)

&lt;223&gt; Conserved domain

&lt;400&gt; 200



Met Met Asn Pro Phe Leu Pro Glu Gly Cys Asp Pro Pro Pro Pro Pro  
 1 5 10 15  
 Gln Pro Met Glu Gly Leu His Glu Asn Ala Pro Pro Pro Phe Leu Thr  
 20 25 30  
 Lys Thr Phe Glu Met Val Asp Asp Pro Asn Thr Asp His Ile Val Ser  
 35 40 45  
 Trp Asn Arg Gly Gly Thr Ser Phe Val Val Trp Asp Leu His Ser Phe  
 50 55 60  
 Ser Thr Ile Leu Leu Pro Arg His Phe Lys His Ser Asn Phe Ser Ser  
 65 70 75 80  
 Phe Ile Arg Gln Leu Asn Thr Tyr Gly Phe Arg Lys Ile Glu Ala Glu  
 85 90 95  
 Arg Trp Glu Phe Ala Asn Glu Glu Phe Leu Leu Gly Gln Arg Gln Leu  
 100 105 110  
 Leu Lys Asn Ile Lys Arg Arg Asn Pro Phe Thr Pro Ser Ser Ser Pro  
 115 120 125  
 Ser His Asp Ala Cys Asn Glu Leu Arg Arg Glu Lys Gln Val Leu Met  
 130 135 140  
 Met Glu Ile Val Ser Leu Arg Gln Gln Gln Gln Thr Thr Lys Ser Tyr  
 145 150 155 160  
 Ile Lys Ala Met Glu Gln Arg Ile Glu Gly Thr Glu Arg Lys Gln Arg  
 165 170 175  
 Gln Met Met Ser Phe Leu Ala Arg Ala Met Gln Ser Pro Ser Phe Leu  
 180 185 190  
 His Gln Leu Leu Lys Gln Arg Asp Lys Lys Ile Lys Glu Leu Glu Asp  
 195 200 205  
 Asn Glu Ser Ala Lys Arg Lys Arg Gly Ser Ser Ser Met Ser Glu Leu  
 210 215 220  
 Glu Val Leu Ala Leu Glu Met Gln Gly His Gly Lys Gln Arg Asn Met  
 225 230 235 240  
 Leu Glu Glu Glu Asp His Gln Leu Val Val Glu Arg Glu Leu Asp Asp  
 245 250 255  
 Gly Phe Trp Glu Glu Leu Leu Ser Asp Glu Ser Leu Ala Ser Thr Ser  
 260 265 270

&lt;210&gt; 201

&lt;211&gt; 1126

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (282)...(920)

&lt;400&gt; 201

atccccactt gttgttcac accaagccaa gtcctatgtc ctagtcactc cacagattcc 60  
 ctatcatcat caattcgttt caaacttagt tcctttcaaa gtcttgtaca tatatacaca 120  
 cacacctatt attctcttgg tgtgtttgtg tgttacatat acgtgtgagt acatactttg 180  
 ttgtaaaagt ggatcggagg tatggaaaagg gaccggttcc accggaaaca tcggcggcgg 240  
 cggatgataa ttcgtcttgg aacgagactg atgtcaccgc c atg gtc tcc gct ctc 296  
 Met Val Ser Ala Leu  
 1 5

agc cgt gtc ata gag aat ccg aca gac ccg ccg gtc aaa caa gag ctt 344  
 Ser Arg Val Ile Glu Asn Pro Thr Asp Pro Pro Val Lys Gln Glu Leu  
 10 15 20

```

gat aaa tcg gat caa cat caa cca gac caa gat caa cca aga aga aga 392
Asp Lys Ser Asp Gln His Gln Pro Asp Gln Asp Gln Pro Arg Arg Arg
      25              30              35

cac tat aga ggc gta agg cag aga cca tgg ggt aaa tgg gcg gca gaa 440
His Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Trp Ala Ala Glu
      40              45              50

atc cgc gat cca aag aaa gca gcc cgt gtc tgg ctc ggg act ttc gag 488
Ile Arg Asp Pro Lys Lys Ala Ala Arg Val Trp Leu Gly Thr Phe Glu
      55              60              65

acg gca gag gaa gct gct tta gcc tat gac cga gct gcc ctc aaa ttc 536
Thr Ala Glu Glu Ala Ala Leu Ala Tyr Asp Arg Ala Ala Leu Lys Phe
      70              75              80              85

aaa ggc acc aag gct aaa ctg aac ttc cct gaa cgg gtc caa ggc cct 584
Lys Gly Thr Lys Ala Lys Leu Asn Phe Pro Glu Arg Val Gln Gly Pro
      90              95              100

act acc acc aca acc att tct cat gca cca aga gga gtt agt gaa tcc 632
Thr Thr Thr Thr Thr Ile Ser His Ala Pro Arg Gly Val Ser Glu Ser
      105              110              115

atg aac tca cct cct cct cga cct ggt cca cct tca act act act act 680
Met Asn Ser Pro Pro Pro Arg Pro Gly Pro Pro Ser Thr Thr Thr Thr
      120              125              130

tcg tgg cca atg act tat aac cag gac ata ctt caa tac gct cag ttg 728
Ser Trp Pro Met Thr Tyr Asn Gln Asp Ile Leu Gln Tyr Ala Gln Leu
      135              140              145

ctt acg agt aac aat gag gtt gat tta tca tac tac acg tcg act ctc 776
Leu Thr Ser Asn Asn Glu Val Asp Leu Ser Tyr Tyr Thr Ser Thr Leu
      150              155              160              165

ttc agt caa cct ttt tca acg cct tct tca tct tct tct tcc tcc caa 824
Phe Ser Gln Pro Phe Ser Thr Pro Ser Ser Ser Ser Ser Ser Gln
      170              175              180

cag acg cag caa cag cag cta caa caa caa caa cag cag cgt gaa gaa 872
Gln Thr Gln Gln Gln Gln Leu Gln Gln Gln Gln Gln Gln Arg Glu Glu
      185              190              195

gaa gag aag aat tat ggt tac aat tat tat aac tac cca aga gaa taa 920
Glu Glu Lys Asn Tyr Gly Tyr Asn Tyr Tyr Asn Tyr Pro Arg Glu *
      200              205              210

tctaattatt attgttggtc gaatcagttt tataaatagc tatcatagtt tcatttttgg 980
tttcgctaac ctttgttgca tggaaaatat gaatgaacga gggacatgtg taacaatttg 1040
tttgtgtttc gtaaattgta gttgtatttg gatttgctga agtttgattt tctgagcata 1100
aatcatttga cgggtcaaaaa aaaaaa 1126

```

&lt;210&gt; 202

&lt;211&gt; 212

&lt;212&gt; PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (36)...(103)

<223> Conserved domain

<400> 202

```

Met Val Ser Ala Leu Ser Arg Val Ile Glu Asn Pro Thr Asp Pro Pro
 1          5          10          15
Val Lys Gln Glu Leu Asp Lys Ser Asp Gln His Gln Pro Asp Gln Asp
 20          25          30
Gln Pro Arg Arg Arg His Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly
 35          40          45
Lys Trp Ala Ala Glu Ile Arg Asp Pro Lys Lys Ala Ala Arg Val Trp
 50          55          60
Leu Gly Thr Phe Glu Thr Ala Glu Glu Ala Ala Leu Ala Tyr Asp Arg
 65          70          75          80
Ala Ala Leu Lys Phe Lys Gly Thr Lys Ala Lys Leu Asn Phe Pro Glu
 85          90          95
Arg Val Gln Gly Pro Thr Thr Thr Thr Thr Ile Ser His Ala Pro Arg
100          105          110
Gly Val Ser Glu Ser Met Asn Ser Pro Pro Pro Arg Pro Gly Pro Pro
115          120          125
Ser Thr Thr Thr Thr Ser Trp Pro Met Thr Tyr Asn Gln Asp Ile Leu
130          135          140
Gln Tyr Ala Gln Leu Leu Thr Ser Asn Asn Glu Val Asp Leu Ser Tyr
145          150          155          160
Tyr Thr Ser Thr Leu Phe Ser Gln Pro Phe Ser Thr Pro Ser Ser Ser
165          170          175
Ser Ser Ser Ser Gln Gln Thr Gln Gln Gln Gln Leu Gln Gln Gln Gln
180          185          190
Gln Gln Arg Glu Glu Glu Glu Lys Asn Tyr Gly Tyr Asn Tyr Tyr Asn
195          200          205
Tyr Pro Arg Glu
210

```

<210> 203

<211> 2057

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (102)...(1805)

<400> 203

```

ttccccctgag aacgacagga gaaagaataa aaaccctaaa tttctttaat ttcggcgctt 60
cagattatcg ttgttaaagg tttttgattg attttgttta a atg ggc gat ctt gct 116
                                     Met Gly Asp Leu Ala
                                     1          5

atg tcc gta gca gac atc agg atg gag aat gag cct gat gat tta gct 164
Met Ser Val Ala Asp Ile Arg Met Glu Asn Glu Pro Asp Asp Leu Ala
      10          15          20

```

agt gat aat gtt gct gag att gat gtg agt gat gaa gag att gat gct	212
Ser Asp Asn Val Ala Glu Ile Asp Val Ser Asp Glu Glu Ile Asp Ala	
25 30 35	
gac gac ctt gag aga cgg atg tgg aaa gat cgt gtc agg ctt aaa aga	260
Asp Asp Leu Glu Arg Arg Met Trp Lys Asp Arg Val Arg Leu Lys Arg	
40 45 50	
atc aaa gag cga caa aaa gct ggc tct caa gga gct caa acg aag gag	308
Ile Lys Glu Arg Gln Lys Ala Gly Ser Gln Gly Ala Gln Thr Lys Glu	
55 60 65	
aca cct aag aaa atc tct gat caa gct cag agg aag aaa atg tct aga	356
Thr Pro Lys Lys Ile Ser Asp Gln Ala Gln Arg Lys Lys Met Ser Arg	
70 75 80 85	
gct caa gat ggt atc ctt aag tac atg ttg aag ctt atg gaa gtc tgc	404
Ala Gln Asp Gly Ile Leu Lys Tyr Met Leu Lys Leu Met Glu Val Cys	
90 95 100	
aaa gtt cgc ggg ttt gtc tat ggt ata ata ccg gaa aag ggc aag cct	452
Lys Val Arg Gly Phe Val Tyr Gly Ile Ile Pro Glu Lys Gly Lys Pro	
105 110 115	
gtg agt ggt tcc tct gac aat ata aga gct tgg tgg aaa gag aaa gtg	500
Val Ser Gly Ser Ser Asp Asn Ile Arg Ala Trp Trp Lys Glu Lys Val	
120 125 130	
aag ttt gat aag aac ggt cct gct gct att gct aaa tac gaa gag gag	548
Lys Phe Asp Lys Asn Gly Pro Ala Ala Ile Ala Lys Tyr Glu Glu Glu	
135 140 145	
tgt tta gcg ttt ggg aaa tct gat ggg aat agg aat tca cag ttt gtt	596
Cys Leu Ala Phe Gly Lys Ser Asp Gly Asn Arg Asn Ser Gln Phe Val	
150 155 160 165	
ctc cag gat ttg caa gat gct act tta ggg tct ttg tta tct tct ttg	644
Leu Gln Asp Leu Gln Asp Ala Thr Leu Gly Ser Leu Leu Ser Ser Leu	
170 175 180	
atg caa cat tgt gat cct cct caa agg aag tat ccg ttg gag aaa ggg	692
Met Gln His Cys Asp Pro Pro Gln Arg Lys Tyr Pro Leu Glu Lys Gly	
185 190 195	
acg cct ccg cct tgg tgg cca acg ggg aat gaa gaa tgg tgg gtg aaa	740
Thr Pro Pro Pro Trp Trp Pro Thr Gly Asn Glu Glu Trp Trp Val Lys	
200 205 210	
ctc ggt ctg cct aaa agc cag agt cct cct tac cga aaa cct cat gat	788
Leu Gly Leu Pro Lys Ser Gln Ser Pro Pro Tyr Arg Lys Pro His Asp	
215 220 225	
ctc aag aag atg tgg aag gtt gga gtt tta acg gca gtg atc aat cat	836
Leu Lys Lys Met Trp Lys Val Gly Val Leu Thr Ala Val Ile Asn His	
230 235 240 245	
atg tta cct gat att gca aag att aag agg cat gtt cgt cag tcg aaa	884

Met Leu Pro Asp	Ile Ala Lys Ile Lys Arg His Val Arg Gln Ser Lys	
	250 255 260	
tgt tta cag gac aag atg aca gct aaa gag agt gcg att tgg ttg gcg	932	
Cys Leu Gln Asp Lys Met Thr Ala Lys Glu Ser Ala Ile Trp Leu Ala	265 270 275	
gtt ttg aac caa gag gaa tct ttg att cag cag cct agc agt gac aat	980	
Val Leu Asn Gln Glu Glu Ser Leu Ile Gln Gln Pro Ser Ser Asp Asn	280 285 290	
gga aac tcc aat gtg act gag aca cat cgt agg ggt aat aac gct gac	1028	
Gly Asn Ser Asn Val Thr Glu Thr His Arg Arg Gly Asn Asn Ala Asp	295 300 305	
agg agg aaa cct gtg gtc aac agt gac agt gac tat gat gtt gat ggg	1076	
Arg Arg Lys Pro Val Val Asn Ser Asp Ser Asp Tyr Asp Val Asp Gly	310 315 320 325	
aca gag gaa gct tca ggt tca gtt tca tct aaa gac agt aga aga aat	1124	
Thr Glu Glu Ala Ser Gly Ser Val Ser Ser Lys Asp Ser Arg Arg Asn	330 335 340	
cag att caa aaa gaa caa cca aca gcc atc tca cat tca gta aga gat	1172	
Gln Ile Gln Lys Glu Gln Pro Thr Ala Ile Ser His Ser Val Arg Asp	345 350 355	
caa gat aaa gca gag aaa cat cgc aga agg aaa aga cct cga att aga	1220	
Gln Asp Lys Ala Glu Lys His Arg Arg Arg Lys Arg Pro Arg Ile Arg	360 365 370	
tcc gga act gtc aat cga caa gag gaa gaa caa cct gaa gct caa caa	1268	
Ser Gly Thr Val Asn Arg Gln Glu Glu Glu Gln Pro Glu Ala Gln Gln	375 380 385	
aga aac atc tta cct gat atg aat cat gtt gat gcc cct ctg cta gaa	1316	
Arg Asn Ile Leu Pro Asp Met Asn His Val Asp Ala Pro Leu Leu Glu	390 395 400 405	
tat aac atc aac ggt act cat caa gag gac gat gtt gtc gac cca aat	1364	
Tyr Asn Ile Asn Gly Thr His Gln Glu Asp Asp Val Val Asp Pro Asn	410 415 420	
att gcc tta gga cca gag gat aat ggt ctg gaa cta gtg gtt cct gag	1412	
Ile Ala Leu Gly Pro Glu Asp Asn Gly Leu Glu Leu Val Val Pro Glu	425 430 435	
ttc aat aac aac tat act tat ctt cca ctt gtt aat gaa caa act atg	1460	
Phe Asn Asn Asn Tyr Thr Tyr Leu Pro Leu Val Asn Glu Gln Thr Met	440 445 450	
atg cct gta gac gaa agg cca atg ctt tat gga cca aac cct aac caa	1508	
Met Pro Val Asp Glu Arg Pro Met Leu Tyr Gly Pro Asn Pro Asn Gln	455 460 465	
gag ctt caa ttt ggg tca ggg tac aac ttc tac aat ccc tct gca gtg	1556	
Glu Leu Gln Phe Gly Ser Gly Tyr Asn Phe Tyr Asn Pro Ser Ala Val		

```

470          475          480          485
ttt gta cat aac cag gaa gac gac att ctc cat aca cag ata gaa atg 1604
Phe Val His Asn Gln Glu Asp Asp Ile Leu His Thr Gln Ile Glu Met
          490          495          500

aat aca caa gca cca cct cac aac agt ggg ttc gag gag gcc cca gga 1652
Asn Thr Gln Ala Pro Pro His Asn Ser Gly Phe Glu Glu Ala Pro Gly
          505          510          515

gga gta ctt caa ccc ctt ggt tta ctc gga aat gaa gac ggt gta aca 1700
Gly Val Leu Gln Pro Leu Gly Leu Leu Gly Asn Glu Asp Gly Val Thr
          520          525          530

ggg agt gag ttg cct cag tat cag agt ggc att ctg tct cca ttg act 1748
Gly Ser Glu Leu Pro Gln Tyr Gln Ser Gly Ile Leu Ser Pro Leu Thr
          535          540          545

gac ttg gac ttt gac tat ggt ggt ttt ggt gat gat ttc tca tgg ttt 1796
Asp Leu Asp Phe Asp Tyr Gly Gly Phe Gly Asp Asp Phe Ser Trp Phe
550          555          560          565

gga gct tag tgtcttgcca ttttttttgg gagattacat agttcaaaag 1845
Gly Ala *

```

```

gacatggcaa tagtctgggt agtacagtta ctttctcttc ttcatttctt ctgatcttat 1905
attcttcttc tttttttctt ataatatattt cttagatttg ttaagagaaa caattttcct 1965
tttgaataag ttgccagaag aactgctttg cccgttgtaa tggctctctag ggaaagcagt 2025
tagcgtatca tcatttgtaa atttactgtg ag 2057

```

<210> 204  
 <211> 567  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (86)...(96)  
 <223> Conserved domain

```

<400> 204
Met Gly Asp Leu Ala Met Ser Val Ala Asp Ile Arg Met Glu Asn Glu
 1          5          10          15
Pro Asp Asp Leu Ala Ser Asp Asn Val Ala Glu Ile Asp Val Ser Asp
          20          25          30
Glu Glu Ile Asp Ala Asp Asp Leu Glu Arg Arg Met Trp Lys Asp Arg
          35          40          45
Val Arg Leu Lys Arg Ile Lys Glu Arg Gln Lys Ala Gly Ser Gln Gly
          50          55          60
Ala Gln Thr Lys Glu Thr Pro Lys Lys Ile Ser Asp Gln Ala Gln Arg
          65          70          75          80
Lys Lys Met Ser Arg Ala Gln Asp Gly Ile Leu Lys Tyr Met Leu Lys
          85          90          95
Leu Met Glu Val Cys Lys Val Arg Gly Phe Val Tyr Gly Ile Ile Pro
          100          105          110
Glu Lys Gly Lys Pro Val Ser Gly Ser Ser Asp Asn Ile Arg Ala Trp

```

	115					120					125				
Trp	Lys	Glu	Lys	Val	Lys	Phe	Asp	Lys	Asn	Gly	Pro	Ala	Ala	Ile	Ala
	130					135					140				
Lys	Tyr	Glu	Glu	Glu	Cys	Leu	Ala	Phe	Gly	Lys	Ser	Asp	Gly	Asn	Arg
145					150					155					160
Asn	Ser	Gln	Phe	Val	Leu	Gln	Asp	Leu	Gln	Asp	Ala	Thr	Leu	Gly	Ser
				165						170				175	
Leu	Leu	Ser	Ser	Leu	Met	Gln	His	Cys	Asp	Pro	Pro	Gln	Arg	Lys	Tyr
			180					185						190	
Pro	Leu	Glu	Lys	Gly	Thr	Pro	Pro	Pro	Trp	Trp	Pro	Thr	Gly	Asn	Glu
	195					200						205			
Glu	Trp	Trp	Val	Lys	Leu	Gly	Leu	Pro	Lys	Ser	Gln	Ser	Pro	Pro	Tyr
	210					215					220				
Arg	Lys	Pro	His	Asp	Leu	Lys	Lys	Met	Trp	Lys	Val	Gly	Val	Leu	Thr
225					230					235					240
Ala	Val	Ile	Asn	His	Met	Leu	Pro	Asp	Ile	Ala	Lys	Ile	Lys	Arg	His
				245					250					255	
Val	Arg	Gln	Ser	Lys	Cys	Leu	Gln	Asp	Lys	Met	Thr	Ala	Lys	Glu	Ser
			260					265					270		
Ala	Ile	Trp	Leu	Ala	Val	Leu	Asn	Gln	Glu	Glu	Ser	Leu	Ile	Gln	Gln
	275						280					285			
Pro	Ser	Ser	Asp	Asn	Gly	Asn	Ser	Asn	Val	Thr	Glu	Thr	His	Arg	Arg
	290					295					300				
Gly	Asn	Asn	Ala	Asp	Arg	Arg	Lys	Pro	Val	Val	Asn	Ser	Asp	Ser	Asp
305				310						315					320
Tyr	Asp	Val	Asp	Gly	Thr	Glu	Glu	Ala	Ser	Gly	Ser	Val	Ser	Ser	Lys
				325					330					335	
Asp	Ser	Arg	Arg	Asn	Gln	Ile	Gln	Lys	Glu	Gln	Pro	Thr	Ala	Ile	Ser
			340					345						350	
His	Ser	Val	Arg	Asp	Gln	Asp	Lys	Ala	Glu	Lys	His	Arg	Arg	Arg	Lys
	355						360					365			
Arg	Pro	Arg	Ile	Arg	Ser	Gly	Thr	Val	Asn	Arg	Gln	Glu	Glu	Glu	Gln
	370					375					380				
Pro	Glu	Ala	Gln	Gln	Arg	Asn	Ile	Leu	Pro	Asp	Met	Asn	His	Val	Asp
385					390					395					400
Ala	Pro	Leu	Leu	Glu	Tyr	Asn	Ile	Asn	Gly	Thr	His	Gln	Glu	Asp	Asp
				405					410					415	
Val	Val	Asp	Pro	Asn	Ile	Ala	Leu	Gly	Pro	Glu	Asp	Asn	Gly	Leu	Glu
			420					425					430		
Leu	Val	Val	Pro	Glu	Phe	Asn	Asn	Asn	Tyr	Thr	Tyr	Leu	Pro	Leu	Val
	435						440					445			
Asn	Glu	Gln	Thr	Met	Met	Pro	Val	Asp	Glu	Arg	Pro	Met	Leu	Tyr	Gly
	450					455					460				
Pro	Asn	Pro	Asn	Gln	Glu	Leu	Gln	Phe	Gly	Ser	Gly	Tyr	Asn	Phe	Tyr
465				470					475						

<210> 205  
 <211> 1638  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (267)...(1259)

<400> 205  
 ggaatttcgg atcgtgtctc tctctgtttc tttgtttcaa tccgatttcg aatcaagccc 60  
 ttactttgtg caccttcaag atttcgtttt ttccagcgcc cagaatgctc cgggtgacca 120  
 acatttggtc ctgattcatt tcctattggt tcgtattgtc tgtgcacaca agagaaattt 180  
 caagaagttg ttactaaaag agaggccaca agtggatatt gtctttgtta tcaagtgtta 240  
 gtacagaaaa gtggtgagaa agtaat atg gct gat acc agt ccg aga act gat 293  
 Met Ala Asp Thr Ser Pro Arg Thr Asp  
 1 5

gct tca aca gat gac gac aca gat cat cct gat ctt ggg tcg gag gga 341  
 Val Ser Thr Asp Asp Asp Thr Asp His Pro Asp Leu Gly Ser Glu Gly  
 10 15 20 25

gca cta gtg aat act gct gct tct gat tcg agt gac cga tcg aag gga 389  
 Ala Leu Val Asn Thr Ala Ala Ser Asp Ser Ser Asp Arg Ser Lys Gly  
 30 35 40

aag atg gat caa aag act ctt cgt agg ctt gct caa aac cgt gag gca 437  
 Lys Met Asp Gln Lys Thr Leu Arg Arg Leu Ala Gln Asn Arg Glu Ala  
 45 50 55

gca agg aaa agc aga ttg agg aag aag gct tat gtt cag cag cta gag 485  
 Ala Arg Lys Ser Arg Leu Arg Lys Lys Ala Tyr Val Gln Gln Leu Glu  
 60 65 70

aac agc cgc ttg aaa cta acc cag ctt gag cag gag ctg caa aga gca 533  
 Asn Ser Arg Leu Lys Leu Thr Gln Leu Glu Gln Glu Leu Gln Arg Ala  
 75 80 85

aga cag cag ggc gtc ttc att tca ggc aca gga gac cag gcc cat tct 581  
 Arg Gln Gln Gly Val Phe Ile Ser Gly Thr Gly Asp Gln Ala His Ser  
 90 95 100 105

act ggt gga aat ggt gct ttg gcg ttt gat gct gaa cat tca cgg tgg 629  
 Thr Gly Gly Asn Gly Ala Leu Ala Phe Asp Ala Glu His Ser Arg Trp  
 110 115 120

ttg gaa gaa aag aac aag caa atg aac gag ctg agg tct gct ctg aat 677  
 Leu Glu Glu Lys Asn Lys Gln Met Asn Glu Leu Arg Ser Ala Leu Asn  
 125 130 135

gcg cat gca ggt gat tct gag ctt cga ata ata gtc gat ggt gtg atg 725  
 Ala His Ala Gly Asp Ser Glu Leu Arg Ile Ile Val Asp Gly Val Met  
 140 145 150

gct cac tat gag gag ctt ttc agg ata aag agc aat gca gct aag aat 773



Ala His Tyr Glu Glu Leu Phe Arg Ile Lys Ser Asn Ala Ala Lys Asn  
 155 160 165

gat gtc ttt cac ttg cta tct ggc atg tgg aaa aca cca gct gag aga 821  
 Asp Val Phe His Leu Leu Ser Gly Met Trp Lys Thr Pro Ala Glu Arg  
 170 175 180 185

tgt ttc ttg tgg ctc ggt gga ttt cgt tca tcc gaa ctt cta aag ctt 869  
 Cys Phe Leu Trp Leu Gly Gly Phe Arg Ser Ser Glu Leu Leu Lys Leu  
 190 195 200

ctg gcg aat cag ttg gag cca atg aca gag aga cag ttg atg ggc ata 917  
 Leu Ala Asn Gln Leu Glu Pro Met Thr Glu Arg Gln Leu Met Gly Ile  
 205 210 215

aat aac ctg caa cag aca tcg cag cag gct gaa gat gct ttg tct caa 965  
 Asn Asn Leu Gln Gln Thr Ser Gln Gln Ala Glu Asp Ala Leu Ser Gln  
 220 225 230

ggg atg gag agc tta caa cag tca cta gct gat act tta tcg agc ggg 1013  
 Gly Met Glu Ser Leu Gln Gln Ser Leu Ala Asp Thr Leu Ser Ser Gly  
 235 240 245

act ctt ggt tca agt tca tca ggg aat gtc gca agc tac atg ggt cag 1061  
 Thr Leu Gly Ser Ser Ser Ser Gly Asn Val Ala Ser Tyr Met Gly Gln  
 250 255 260 265

atg gcc atg gca atg gga aag tta ggt aca ctc gaa gga ttt atc cgc 1109  
 Met Ala Met Ala Met Gly Lys Leu Gly Thr Leu Glu Gly Phe Ile Arg  
 270 275 280

cag gct gat aat ttg aga cta caa aca ttg caa cag atg ata aga gta 1157  
 Gln Ala Asp Asn Leu Arg Leu Gln Thr Leu Gln Gln Met Ile Arg Val  
 285 290 295

tta aca acg aga cag tca gca cgt gct cta ctt gca ata cac gat tac 1205  
 Leu Thr Thr Arg Gln Ser Ala Arg Ala Leu Leu Ala Ile His Asp Tyr  
 300 305 310

ttc tca cgg cta cga gct cta agc tcc tta tgg ctt gct cga ccc aga 1253  
 Phe Ser Arg Leu Arg Ala Leu Ser Ser Leu Trp Leu Ala Arg Pro Arg  
 315 320 325

gag tga aactgtatatt tggtcacatg tcagctgtac aaaatccata tggacacaaa 1309  
 Glu \*  
 330

accaggagag actathtaate aacacttgctc agattcttct taccaaatec atcaacaaat 1369  
 aagcaaatct ctgggaaaca aaagactctt tgatgtagg tttcttctac atgggtgtgg 1429  
 taattcatgt tgttttagtt gtagtcatca gtttttaatt tagcatttga aaagttcaat 1489  
 gttgtttata tagcatcttc gattatctta gaaagggttat tgaattttgt tttttttgt 1549  
 tacttttgtg tgtggtaaag gtgttttaac cttgcaactt ctgtactgta atcatttaac 1609  
 aatattaaga tgttctattt gagttttgt 1638

&lt;210&gt; 206

&lt;211&gt; 330

&lt;212&gt; PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (45)...(105)

<223> Conserved domain

<400> 206

```

Met Ala Asp Thr Ser Pro Arg Thr Asp Val Ser Thr Asp Asp Asp Thr
 1           5           10           15
Asp His Pro Asp Leu Gly Ser Glu Gly Ala Leu Val Asn Thr Ala Ala
      20           25           30
Ser Asp Ser Ser Asp Arg Ser Lys Gly Lys Met Asp Gln Lys Thr Leu
 35           40           45
Arg Arg Leu Ala Gln Asn Arg Glu Ala Ala Arg Lys Ser Arg Leu Arg
 50           55           60
Lys Lys Ala Tyr Val Gln Gln Leu Glu Asn Ser Arg Leu Lys Leu Thr
 65           70           75           80
Gln Leu Glu Gln Glu Leu Gln Arg Ala Arg Gln Gln Gly Val Phe Ile
      85           90           95
Ser Gly Thr Gly Asp Gln Ala His Ser Thr Gly Gly Asn Gly Ala Leu
      100           105           110
Ala Phe Asp Ala Glu His Ser Arg Trp Leu Glu Glu Lys Asn Lys Gln
      115           120           125
Met Asn Glu Leu Arg Ser Ala Leu Asn Ala His Ala Gly Asp Ser Glu
 130           135           140
Leu Arg Ile Ile Val Asp Gly Val Met Ala His Tyr Glu Glu Leu Phe
 145           150           155           160
Arg Ile Lys Ser Asn Ala Ala Lys Asn Asp Val Phe His Leu Leu Ser
      165           170           175
Gly Met Trp Lys Thr Pro Ala Glu Arg Cys Phe Leu Trp Leu Gly Gly
      180           185           190
Phe Arg Ser Ser Glu Leu Leu Lys Leu Leu Ala Asn Gln Leu Glu Pro
      195           200           205
Met Thr Glu Arg Gln Leu Met Gly Ile Asn Asn Leu Gln Gln Thr Ser
 210           215           220
Gln Gln Ala Glu Asp Ala Leu Ser Gln Gly Met Glu Ser Leu Gln Gln
 225           230           235           240
Ser Leu Ala Asp Thr Leu Ser Ser Gly Thr Leu Gly Ser Ser Ser Ser
      245           250           255
Gly Asn Val Ala Ser Tyr Met Gly Gln Met Ala Met Ala Met Gly Lys
      260           265           270
Leu Gly Thr Leu Glu Gly Phe Ile Arg Gln Ala Asp Asn Leu Arg Leu
 275           280           285
Gln Thr Leu Gln Gln Met Ile Arg Val Leu Thr Thr Arg Gln Ser Ala
 290           295           300
Arg Ala Leu Leu Ala Ile His Asp Tyr Phe Ser Arg Leu Arg Ala Leu
 305           310           315           320
Ser Ser Leu Trp Leu Ala Arg Pro Arg Glu
      325           330

```

<210> 207

<211> 1370

<212> DNA

<213> Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (184) ... (969)

&lt;400&gt; 207

```

gtcgacccac gcgtccgggt ttttctttta tcctcttata gctaactctgg agctctatat 60
agactataaa gggtttttga ttgattcggg agctcgagat ttgacttctt ttagctgatt 120
cggcaagttt gtatctagaa aggatcgatt ggtgaggtca atagtggttg gtgggtttta 180
gta atg gaa gac ggt gag ctt gat ttc tcc aat cag gaa gtg ttt tcg 228
Met Glu Asp Gly Glu Leu Asp Phe Ser Asn Gln Glu Val Phe Ser
      1              5              10              15

agt tcg gag atg ggt gaa tta cca cct agc aat tgt tcg atg gat agt 276
Ser Ser Glu Met Gly Glu Leu Pro Pro Ser Asn Cys Ser Met Asp Ser
              20              25              30

ttc ttt gat ggg ctt tta atg gat act aat gct gct tgt acc cac act 324
Phe Phe Asp Gly Leu Leu Met Asp Thr Asn Ala Ala Cys Thr His Thr
              35              40              45

cac acc tgt aac ccc act gga cca gag aac act cat act cac acg tgc 372
His Thr Cys Asn Pro Thr Gly Pro Glu Asn Thr His Thr His Thr Cys
              50              55              60

ttc cat gtc cac acc aag att ctc ccg gat gag agc gat gaa aaa gtt 420
Phe His Val His Thr Lys Ile Leu Pro Asp Glu Ser Asp Glu Lys Val
              65              70              75

tct act gat gat aca gct gag tct tgt ggg aag aag ggt gaa aag aga 468
Ser Thr Asp Asp Thr Ala Glu Ser Cys Gly Lys Lys Gly Glu Lys Arg
              80              85              90              95

cct ttg gga aac cgg gaa gcg gtt aga aag tat aga gag aag aag aag 516
Pro Leu Gly Asn Arg Glu Ala Val Arg Lys Tyr Arg Glu Lys Lys Lys
              100              105              110

gct aaa gct gct tct ttg gag gat gag gtt gca agg ctt agg gcg gtg 564
Ala Lys Ala Ala Ser Leu Glu Asp Glu Val Ala Arg Leu Arg Ala Val
              115              120              125

aat cag cag ctg gtg aag agg ttg caa aat cag gct acc ttg gaa gct 612
Asn Gln Gln Leu Val Lys Arg Leu Gln Asn Gln Ala Thr Leu Glu Ala
              130              135              140

gag gtt tcg agg ctt aag tgt ttg ctt gtg gat ttg aga gga aga ata 660
Glu Val Ser Arg Leu Lys Cys Leu Leu Val Asp Leu Arg Gly Arg Ile
              145              150              155

gat gga gag att gga tct ttt cct tat cag aaa cct atg gct gca aat 708
Asp Gly Glu Ile Gly Ser Phe Pro Tyr Gln Lys Pro Met Ala Ala Asn
              160              165              170              175

att cct tct ttc tcg cac atg atg aat cct tgt aat gta caa tgt gat 756
Ile Pro Ser Phe Ser His Met Met Asn Pro Cys Asn Val Gln Cys Asp
              180              185              190

gat gaa gtt tat tgc cct cag aat gtg ttt gga gtg aat agc caa gaa 804

```

Asp Glu Val Tyr Cys Pro Gln Asn Val Phe Gly Val Asn Ser Gln Glu  
 195 200 205

ggg gcc tcg atc aat gac caa ggg tta agt ggt tgt gat ttt gat cag 852  
 Gly Ala Ser Ile Asn Asp Gln Gly Leu Ser Gly Cys Asp Phe Asp Gln  
 210 215 220

cta caa tgc atg gct aat cag aac tta aat gga aat gga aac gga tca 900  
 Leu Gln Cys Met Ala Asn Gln Asn Leu Asn Gly Asn Gly Asn Gly Ser  
 225 230 235

ttc agc aac gtc aat aca tct gtc tcg aat aag aga aaa ggt ggg cat 948  
 Phe Ser Asn Val Asn Thr Ser Val Ser Asn Lys Arg Lys Gly Gly His  
 240 245 250 255

cgt gca tca aga gca gtt tga agcatcatca agcttgact atctatttcc 999  
 Arg Ala Ser Arg Ala Val \*  
 260

accagcatag atattgtatt ccaaataagt ttagaggttc agctgcagga tcagcttcgc 1059  
 tcagctttga ggggttggtg gtgtggtcctt tctttgtggc acgagtgaga tctatggaca 1119  
 gaaccagat ttagtagtag tagaggcagg atttcgactt ccactaacca tcatgttgct 1179  
 tggatgaagaa caaggtatgc ccatgaagca cactgttttg tacattgagc ttgaggggct 1239  
 gtctctgac tagccttact gtaacattgc aacgttctca caattgtgat cccaagttgc 1299  
 tttgttgact taaatgtgat aatatagctt aacttttact tgaaaaaaaa aaaaaaaaaa 1359  
 aaaaaaaaaa a 1370

<210> 208  
 <211> 261  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (90)...(153)  
 <223> Conserved domain

<400> 208  
 Met Glu Asp Gly Glu Leu Asp Phe Ser Asn Gln Glu Val Phe Ser Ser  
 1 5 10 15  
 Ser Glu Met Gly Glu Leu Pro Pro Ser Asn Cys Ser Met Asp Ser Phe  
 20 25 30  
 Phe Asp Gly Leu Leu Met Asp Thr Asn Ala Ala Cys Thr His Thr His  
 35 40 45  
 Thr Cys Asn Pro Thr Gly Pro Glu Asn Thr His Thr His Thr Cys Phe  
 50 55 60  
 His Val His Thr Lys Ile Leu Pro Asp Glu Ser Asp Glu Lys Val Ser  
 65 70 75 80  
 Thr Asp Asp Thr Ala Glu Ser Cys Gly Lys Lys Gly Glu Lys Arg Pro  
 85 90 95  
 Leu Gly Asn Arg Glu Ala Val Arg Lys Tyr Arg Glu Lys Lys Lys Ala  
 100 105 110  
 Lys Ala Ala Ser Leu Glu Asp Glu Val Ala Arg Leu Arg Ala Val Asn  
 115 120 125  
 Gln Gln Leu Val Lys Arg Leu Gln Asn Gln Ala Thr Leu Glu Ala Glu  
 130 135 140  
 Val Ser Arg Leu Lys Cys Leu Leu Val Asp Leu Arg Gly Arg Ile Asp

145                      150                      155                      160  
 Gly Glu Ile Gly Ser Phe Pro Tyr Gln Lys Pro Met Ala Ala Asn Ile  
                                  165                      170                      175  
 Pro Ser Phe Ser His Met Met Asn Pro Cys Asn Val Gln Cys Asp Asp  
                                  180                      185                      190  
 Glu Val Tyr Cys Pro Gln Asn Val Phe Gly Val Asn Ser Gln Glu Gly  
                                  195                      200                      205  
 Ala Ser Ile Asn Asp Gln Gly Leu Ser Gly Cys Asp Phe Asp Gln Leu  
                                  210                      215                      220  
 Gln Cys Met Ala Asn Gln Asn Leu Asn Gly Asn Gly Asn Gly Ser Phe  
 225                      230                      235                      240  
 Ser Asn Val Asn Thr Ser Val Ser Asn Lys Arg Lys Gly Gly His Arg  
                                  245                      250                      255  
 Ala Ser Arg Ala Val  
                                  260

&lt;210&gt; 209

&lt;211&gt; 859

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (62)...(718)

&lt;400&gt; 209

caatccacta acgatcccta accgaaaaca gagtagtcaa gaaacagagt atttttttcta 60  
 c atg gat cca ttt tta att cag tcc cca ttc tcc ggc ttc tca ccg gaa 109  
   Met Asp Pro Phe Leu Ile Gln Ser Pro Phe Ser Gly Phe Ser Pro Glu  
       1                                  5                                  10                                  15

tat tct atc gga tct tct cca gat tct ttc tca tcc tct tct tct aac 157  
 Tyr Ser Ile Gly Ser Ser Pro Asp Ser Phe Ser Ser Ser Ser Ser Asn  
                                   20                                  25                                  30

aat tac tct ctt ccc ttc aac gag aac gac tca gag gaa atg ttt ctc 205  
 Asn Tyr Ser Leu Pro Phe Asn Glu Asn Asp Ser Glu Glu Met Phe Leu  
                                   35                                  40                                  45

tac ggt cta atc gag cag tcc acg caa caa acc tat att gac tcg gat 253  
 Tyr Gly Leu Ile Glu Gln Ser Thr Gln Gln Thr Tyr Ile Asp Ser Asp  
                                   50                                  55                                  60

agt caa gac ctt ccg atc aaa tcc gta agc tca aga aag tca gag aag 301  
 Ser Gln Asp Leu Pro Ile Lys Ser Val Ser Ser Arg Lys Ser Glu Lys  
                                   65                                  70                                  75                                  80

tct tac aga ggc gta aga cga cgg cca tgg ggg aaa ttc gcg gcg gag 349  
 Ser Tyr Arg Gly Val Arg Arg Arg Pro Trp Gly Lys Phe Ala Ala Glu  
                                   85                                  90                                  95

ata aga gat tcg act aga aac ggt att agg gtt tgg ctc ggg acg ttc 397  
 Ile Arg Asp Ser Thr Arg Asn Gly Ile Arg Val Trp Leu Gly Thr Phe  
                                   100                                  105                                  110

gaa agc gcg gaa gag gcg gct tta gcc tac gat caa gct gct ttc tcg 445

Glu Ser Ala Glu Glu Ala Ala Leu Ala Tyr Asp Gln Ala Ala Phe Ser  
 115 120 125  
 atg aga ggg tcc tcg gcg att ctc aat ttt tcg gcg gag aga gtt caa 493  
 Met Arg Gly Ser Ser Ala Ile Leu Asn Phe Ser Ala Glu Arg Val Gln  
 130 135 140  
 gag tcg ctt tcg gag att aaa tat acc tac gag gat ggt tgt tct ccg 541  
 Glu Ser Leu Ser Glu Ile Lys Tyr Thr Tyr Glu Asp Gly Cys Ser Pro  
 145 150 155 160  
 gtt gtg gcg ttg aag agg aaa cac tcg atg aga cgg aga atg acc aat 589  
 Val Val Ala Leu Lys Arg Lys His Ser Met Arg Arg Arg Met Thr Asn  
 165 170 175  
 aag aag acg aaa gat agt gac ttt gat cac cgc tcc gtg aag tta gat 637  
 Lys Lys Thr Lys Asp Ser Asp Phe Asp His Arg Ser Val Lys Leu Asp  
 180 185 190  
 aat gta gtt gtc ttt gag gat ttg gga gaa cag tac ctt gag gag ctt 685  
 Asn Val Val Val Phe Glu Asp Leu Gly Glu Gln Tyr Leu Glu Glu Leu  
 195 200 205  
 ttg ggg tct tct gaa aat agt ggg act tgg tga aagattagga tttgtattag 738  
 Leu Gly Ser Ser Glu Asn Ser Gly Thr Trp \*  
 210 215  
 ggaccttaag tttgaagtgg ttgattaatt ttaaccctaa tatgtttttt gtttgcttaa 798  
 atatttgatt ctattgagaa acatcgaaaa cagtttgtat gtacttttgt gatacttggc 858  
 g 859

&lt;210&gt; 210

&lt;211&gt; 218

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (79)...(147)

&lt;223&gt; Conserved domain

&lt;400&gt; 210

Met Asp Pro Phe Leu Ile Gln Ser Pro Phe Ser Gly Phe Ser Pro Glu  
 1 5 10 15  
 Tyr Ser Ile Gly Ser Ser Pro Asp Ser Phe Ser Ser Ser Ser Ser Asn  
 20 25 30  
 Asn Tyr Ser Leu Pro Phe Asn Glu Asn Asp Ser Glu Glu Met Phe Leu  
 35 40 45  
 Tyr Gly Leu Ile Glu Gln Ser Thr Gln Gln Thr Tyr Ile Asp Ser Asp  
 50 55 60  
 Ser Gln Asp Leu Pro Ile Lys Ser Val Ser Ser Arg Lys Ser Glu Lys  
 65 70 75 80  
 Ser Tyr Arg Gly Val Arg Arg Arg Pro Trp Gly Lys Phe Ala Ala Glu  
 85 90 95  
 Ile Arg Asp Ser Thr Arg Asn Gly Ile Arg Val Trp Leu Gly Thr Phe  
 100 105 110  
 Glu Ser Ala Glu Glu Ala Ala Leu Ala Tyr Asp Gln Ala Ala Phe Ser

```

      115              120              125
Met Arg Gly Ser Ser Ala Ile Leu Asn Phe Ser Ala Glu Arg Val Gln
      130              135              140
Glu Ser Leu Ser Glu Ile Lys Tyr Thr Tyr Glu Asp Gly Cys Ser Pro
      145              150              155              160
Val Val Ala Leu Lys Arg Lys His Ser Met Arg Arg Arg Met Thr Asn
      165              170              175
Lys Lys Thr Lys Asp Ser Asp Phe Asp His Arg Ser Val Lys Leu Asp
      180              185              190
Asn Val Val Val Phe Glu Asp Leu Gly Glu Gln Tyr Leu Glu Glu Leu
      195              200              205
Leu Gly Ser Ser Glu Asn Ser Gly Thr Trp
      210              215

```

<210> 211  
 <211> 1055  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (70)...(816)

<400> 211  
 ataaaggcat ttcagctcca ccgtaggaaa ctttctcttg aaagaaaccc acagcaacaa 60  
 acagagaaa atg tgt ggc ggt gct att att tcc gat tat gcc cct ctc gtc 111  
 Met Cys Gly Gly Ala Ile Ile Ser Asp Tyr Ala Pro Leu Val  
 1 5 10

acc aag gcc aag ggc cgt aaa ctc acg gct gag gaa ctc tgg tca gag 159  
 Thr Lys Ala Lys Gly Arg Lys Leu Thr Ala Glu Glu Leu Trp Ser Glu  
 15 20 25 30

ctc gat gct tcc gcc gcc gac gac ttc tgg ggt ttc tat tcc acc tcc 207  
 Leu Asp Ala Ser Ala Ala Asp Asp Phe Trp Gly Phe Tyr Ser Thr Ser  
 35 40 45

aaa ctc cat ccc acc aac caa gtt aac gtg aaa gag gag gca gtg aag 255  
 Lys Leu His Pro Thr Asn Gln Val Asn Val Lys Glu Glu Ala Val Lys  
 50 55 60

aag gag cag gca aca gag ccg ggg aaa cgg agg aag agg aag aat gtt 303  
 Lys Glu Gln Ala Thr Glu Pro Gly Lys Arg Arg Lys Arg Lys Asn Val  
 65 70 75

tat aga ggg ata cgt aag cgt cca tgg gga aaa tgg gcg gct gag att 351  
 Tyr Arg Gly Ile Arg Lys Arg Pro Trp Gly Lys Trp Ala Ala Glu Ile  
 80 85 90

cga gat cca cga aaa ggt gtt aga gtt tgg ctt ggt acg ttc aac acg 399  
 Arg Asp Pro Arg Lys Gly Val Arg Val Trp Leu Gly Thr Phe Asn Thr  
 95 100 105 110

gcg gag gaa gct gcc atg gct tat gat gtt gcg gcc aag cag atc cgt 447  
 Ala Glu Glu Ala Ala Met Ala Tyr Asp Val Ala Ala Lys Gln Ile Arg  
 115 120 125

```

ggt gat aaa gcc aag ctc aac ttc cca gat ctg cac cat cct cct cct 495
Gly Asp Lys Ala Lys Leu Asn Phe Pro Asp Leu His His Pro Pro Pro
      130              135              140

cct aat tat act cct ccg ccg tca tcg cca cga tca acc gat cag cct 543
Pro Asn Tyr Thr Pro Pro Pro Ser Ser Pro Arg Ser Thr Asp Gln Pro
      145              150              155

ccg gcg aag aag gtc tgc gtt gtc tct cag agt gag agc gag tta agt 591
Pro Ala Lys Lys Val Cys Val Val Ser Gln Ser Glu Ser Glu Leu Ser
      160              165              170

cag ccg agt ttc ccg gtg gag tgt ata gga ttt gga aat ggg gac gag 639
Gln Pro Ser Phe Pro Val Glu Cys Ile Gly Phe Gly Asn Gly Asp Glu
      175              180              185              190

ttt cag aac ctg agt tac gga ttt gag ccg gat tat gat ctg aaa cag 687
Phe Gln Asn Leu Ser Tyr Gly Phe Glu Pro Asp Tyr Asp Leu Lys Gln
      195              200              205

cag ata tcg agc ttg gaa tcg ttc ctt gag ctg gac ggt aac acg gcg 735
Gln Ile Ser Ser Leu Glu Ser Phe Leu Glu Leu Asp Gly Asn Thr Ala
      210              215              220

gag caa ccg agt cag ctt gat gag tcc gtt tcc gag gtg gat atg tgg 783
Glu Gln Pro Ser Gln Leu Asp Glu Ser Val Ser Glu Val Asp Met Trp
      225              230              235

atg ctt gat gat gtc att gcg tcg tat gag taa aagaaaaaaaa ataagtttaa 836
Met Leu Asp Asp Val Ile Ala Ser Tyr Glu *
      240              245

aaaaagttaa ataaagtctg taatatatat gtaaccgccg ttacttttaa aagggttttta 896
ccgtcgcatg ggactgctga tgatgtctgt tgtgtaatgt gtagaatgtg accaaatgga 956
cggtatatta cggtttggtg tattattagt ttcttagatg gaaaaactta catgtgtaaa 1016
taagatttgt aatgtaagac gaagtactta taacttctt 1055

```

&lt;210&gt; 212

&lt;211&gt; 248

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (76)...(145)

&lt;223&gt; Conserved domain

&lt;400&gt; 212

```

Met Cys Gly Gly-Ala Ile Ile Ser Asp Tyr Ala Pro Leu Val Thr Lys
 1          5          10          15
Ala Lys Gly Arg Lys Leu Thr Ala Glu Glu Leu Trp Ser Glu Leu Asp
      20          25          30
Ala Ser Ala Ala Asp Asp Phe Trp Gly Phe Tyr Ser Thr Ser Lys Leu
      35          40          45
His Pro Thr Asn Gln Val Asn Val Lys Glu Glu Ala Val Lys Lys Glu
      50          55          60

```



Gln Ala Thr Glu Pro Gly Lys Arg Arg Lys Arg Lys Asn Val Tyr Arg  
 65 70 75 80  
 Gly Ile Arg Lys Arg Pro Trp Gly Lys Trp Ala Ala Glu Ile Arg Asp  
 85 90 95  
 Pro Arg Lys Gly Val Arg Val Trp Leu Gly Thr Phe Asn Thr Ala Glu  
 100 105 110  
 Glu Ala Ala Met Ala Tyr Asp Val Ala Ala Lys Gln Ile Arg Gly Asp  
 115 120 125  
 Lys Ala Lys Leu Asn Phe Pro Asp Leu His His Pro Pro Pro Asn  
 130 135 140  
 Tyr Thr Pro Pro Pro Ser Ser Pro Arg Ser Thr Asp Gln Pro Pro Ala  
 145 150 155 160  
 Lys Lys Val Cys Val Val Ser Gln Ser Glu Ser Glu Leu Ser Gln Pro  
 165 170 175  
 Ser Phe Pro Val Glu Cys Ile Gly Phe Gly Asn Gly Asp Glu Phe Gln  
 180 185 190  
 Asn Leu Ser Tyr Gly Phe Glu Pro Asp Tyr Asp Leu Lys Gln Gln Ile  
 195 200 205  
 Ser Ser Leu Glu Ser Phe Leu Glu Leu Asp Gly Asn Thr Ala Glu Gln  
 210 215 220  
 Pro Ser Gln Leu Asp Glu Ser Val Ser Glu Val Asp Met Trp Met Leu  
 225 230 235 240  
 Asp Asp Val Ile Ala Ser Tyr Glu  
 245

&lt;210&gt; 213

&lt;211&gt; 920

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(852)

&lt;400&gt; 213

atg gcg aag acg aaa tat gga gag aga cat agg aaa ggg tta tgg tca 48  
 Met Ala Lys Thr Lys Tyr Gly Glu Arg His Arg Lys Gly Leu Trp Ser  
 1 5 10 15  
 cct gaa gaa gac gag aag cta agg agc ttc atc ctc tct tat ggc cat 96  
 Pro Glu Glu Asp Glu Lys Leu Arg Ser Phe Ile Leu Ser Tyr Gly His  
 20 25 30  
 tct tgc tgg acc act gtt ccc atc aaa gct ggg tta caa agg aat ggg 144  
 Ser Cys Trp Thr Thr Val Pro Ile Lys Ala Gly Leu Gln Arg Asn Gly  
 35 40 45  
 aag agc tgc aga tta aga tgg att aat tac cta aga cca ggg tta aag 192  
 Lys Ser Cys Arg Leu Arg Trp Ile Asn Tyr Leu Arg Pro Gly Leu Lys  
 50 55 60  
 agg gat atg att agt gca gaa gaa gaa gag act atc ttg acg ttt cat 240  
 Arg Asp Met Ile Ser Ala Glu Glu Glu Glu Thr Ile Leu Thr Phe His  
 65 70 75 80  
 tct ccc ttg ggt aac aag tgg tcg caa ata gct aaa ttc tta ccg gga 288

```

Ser Pro Leu Gly Asn Lys Trp Ser Gln Ile Ala Lys Phe Leu Pro Gly
      85                      90                      95

aga aca gac aat gag ata aag aac tat tgg cac tct cat ttg aaa aag   336
Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp His Ser His Leu Lys Lys
      100                      105                      110

aaa tgg ctc aag tct cag agc tta caa gat gca aaa tct att tcc cct   384
Lys Trp Leu Lys Ser Gln Ser Leu Gln Asp Ala Lys Ser Ile Ser Pro
      115                      120                      125

cct tcg tct tca tca tca tca ctt gtt gct tgt gga gaa aga aat ccg   432
Pro Ser Ser Ser Ser Ser Ser Leu Val Ala Cys Gly Glu Arg Asn Pro
      130                      135                      140

gaa acc ttg atc tcg aat cac gtg ttc tcc ctc cag aga ctt cta gag   480
Glu Thr Leu Ile Ser Asn His Val Phe Ser Leu Gln Arg Leu Leu Glu
      145                      150                      155                      160

aac aaa tct tca tct ccc tca caa gaa agc aac gga aat aac agc cat   528
Asn Lys Ser Ser Ser Pro Ser Gln Glu Ser Asn Gly Asn Asn Ser His
      165                      170                      175

caa tgt tct tct gct cct gag att cca agg ctt ttc ttc tct gaa tgg   576
Gln Cys Ser Ser Ala Pro Glu Ile Pro Arg Leu Phe Phe Ser Glu Trp
      180                      185                      190

ctt tct tct tca tat ccc cac acc gat tat tcc tct gag ttt acc gac   624
Leu Ser Ser Ser Tyr Pro His Thr Asp Tyr Ser Ser Glu Phe Thr Asp
      195                      200                      205

tct aag cac agt caa gct cca aat gtc gaa gag act ctc tca gct tat   672
Ser Lys His Ser Gln Ala Pro Asn Val Glu Glu Thr Leu Ser Ala Tyr
      210                      215                      220

gaa gaa atg ggt gat gtt gat cag ttc cat tac aac gaa atg atg atc   720
Glu Glu Met Gly Asp Val Asp Gln Phe His Tyr Asn Glu Met Met Ile
      225                      230                      235                      240

aac aac agc aac tgg act ctt aac gac att gtg ttt ggt tcc aaa tgt   768
Asn Asn Ser Asn Trp Thr Leu Asn Asp Ile Val Phe Gly Ser Lys Cys
      245                      250                      255

aag aag cag gag cat cat att tat aga gag gct tca gat tgt aat tct   816
Lys Lys Gln Glu His His Ile Tyr Arg Glu Ala Ser Asp Cys Asn Ser
      260                      265                      270

tct gct gaa ttc ttt tct cca cca aca acg acg taa attgcgttta       862
Ser Ala Glu Phe Phe Ser Pro Pro Thr Thr Thr *
      275                      280

ttgtaatgta aatcaaattt ctaaggcaaa accggaaaaa aaaaaaaaaa aaaaaaaa   920

```

<210> 214  
 <211> 283  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (11)...(113)  
 <223> Conserved domain

<400> 214  
 Met Ala Lys Thr Lys Tyr Gly Glu Arg His Arg Lys Gly Leu Trp Ser  
 1 5 10 15  
 Pro Glu Glu Asp Glu Lys Leu Arg Ser Phe Ile Leu Ser Tyr Gly His  
 20 25 30  
 Ser Cys Trp Thr Thr Val Pro Ile Lys Ala Gly Leu Gln Arg Asn Gly  
 35 40 45  
 Lys Ser Cys Arg Leu Arg Trp Ile Asn Tyr Leu Arg Pro Gly Leu Lys  
 50 55 60  
 Arg Asp Met Ile Ser Ala Glu Glu Glu Glu Thr Ile Leu Thr Phe His  
 65 70 75 80  
 Ser Pro Leu Gly Asn Lys Trp Ser Gln Ile Ala Lys Phe Leu Pro Gly  
 85 90 95  
 Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp His Ser His Leu Lys Lys  
 100 105 110  
 Lys Trp Leu Lys Ser Gln Ser Leu Gln Asp Ala Lys Ser Ile Ser Pro  
 115 120 125  
 Pro Ser Ser Ser Ser Ser Ser Leu Val Ala Cys Gly Glu Arg Asn Pro  
 130 135 140  
 Glu Thr Leu Ile Ser Asn His Val Phe Ser Leu Gln Arg Leu Leu Glu  
 145 150 155 160  
 Asn Lys Ser Ser Ser Pro Ser Gln Glu Ser Asn Gly Asn Asn Ser His  
 165 170 175  
 Gln Cys Ser Ser Ala Pro Glu Ile Pro Arg Leu Phe Phe Ser Glu Trp  
 180 185 190  
 Leu Ser Ser Ser Tyr Pro His Thr Asp Tyr Ser Ser Glu Phe Thr Asp  
 195 200 205  
 Ser Lys His Ser Gln Ala Pro Asn Val Glu Glu Thr Leu Ser Ala Tyr  
 210 215 220  
 Glu Glu Met Gly Asp Val Asp Gln Phe His Tyr Asn Glu Met Met Ile  
 225 230 235 240  
 Asn Asn Ser Asn Trp Thr Leu Asn Asp Ile Val Phe Gly Ser Lys Cys  
 245 250 255  
 Lys Lys Gln Glu His His Ile Tyr Arg Glu Ala Ser Asp Cys Asn Ser  
 260 265 270  
 Ser Ala Glu Phe Phe Ser Pro Pro Thr Thr Thr  
 275 280

<210> 215  
 <211> 726  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)...(726)

<400> 215  
 atg gcc tgc tca tca tca tca tct tat aga ttc caa tct ggg tct tac 48  
 Met Ala Ser Ser Ser Ser Ser Ser Tyr Arg Phe Gln Ser Gly Ser Tyr

1	5	10	15	
cct ctt tcg tca agt cct tct ctt ggg aat ttc gtc gaa cgc att aaa	96			
Pro Leu Ser Ser Ser Pro Ser Leu Gly Asn Phe Val Glu Arg Ile Lys				
20	25	30		
gac gct tgt cat ttc ctt gtc tct gct gtt ttg ggt acc att atc tcc	144			
Asp Ala Cys His Phe Leu Val Ser Ala Val Leu Gly Thr Ile Ile Ser				
35	40	45		
gcg atc ttg acc ttc ttc ttc gca cta gtg ggc aca ttg cta ggg gca	192			
Ala Ile Leu Thr Phe Phe Phe Ala Leu Val Gly Thr Leu Leu Gly Ala				
50	55	60		
ctt aca gga gct ttg ata ggt caa gaa act gag agt ggt ttc att aga	240			
Leu Thr Gly Ala Leu Ile Gly Gln Glu Thr Glu Ser Gly Phe Ile Arg				
65	70	75	80	
gga gca gca att gga gcc att tcg gga gct gtt ttc tct atc gag gtc	288			
Gly Ala Ala Ile Gly Ala Ile Ser Gly Ala Val Phe Ser Ile Glu Val				
85	90	95		
ttt gaa tca tct ctg gat ctc tgg aaa tcc gat gag tcg ggt ttc gga	336			
Phe Glu Ser Ser Leu Asp Leu Trp Lys Ser Asp Glu Ser Gly Phe Gly				
100	105	110		
tgt ttt ctc tac ttg att gat gtc att gtt agt ctt cta agc ggg aga	384			
Cys Phe Leu Tyr Leu Ile Asp Val Ile Val Ser Leu Leu Ser Gly Arg				
115	120	125		
ctt gta cga gag cgc att ggt cct gca atg cta agt gca gtg caa agt	432			
Leu Val Arg Glu Arg Ile Gly Pro Ala Met Leu Ser Ala Val Gln Ser				
130	135	140		
caa atg gga gct gtg gat aca gct ttt gat gat cac aca agc ctt ttt	480			
Gln Met Gly Ala Val Asp Thr Ala Phe Asp Asp His Thr Ser Leu Phe				
145	150	155	160	
gat aca gga ggc tca aaa gga ttg aca gga gac ctt gtt gag aaa atc	528			
Asp Thr Gly Gly Ser Lys Gly Leu Thr Gly Asp Leu Val Glu Lys Ile				
165	170	175		
cca aag atg aca atc act ggc aac aat aac act gat gct tct gag aac	576			
Pro Lys Met Thr Ile Thr Gly Asn Asn Asn Thr Asp Ala Ser Glu Asn				
180	185	190		
aca gac tca tgt tct gtt tgt ctt cag gat ttc cag ctc ggt gaa aca	624			
Thr Asp Ser Cys Ser Val Cys Leu Gln Asp Phe Gln Leu Gly Glu Thr				
195	200	205		
gtt aga agc ttg cct cat tgt cat cac atg ttt cac tta cct tgc ata	672			
Val Arg Ser Leu Pro His Cys His His Met Phe His Leu Pro Cys Ile				
210	215	220		
gac aat tgg ctc ctt aga cac ggt tct tgc ccg atg tgt aga cgt gat	720			
Asp Asn Trp Leu Leu Arg His Gly Ser Cys Pro Met Cys Arg Arg Asp				
225	230	235	240	

att taa  
Ile \*

726

<210> 216  
<211> 241  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<221> DOMAIN  
<222> (196)...(237)  
<223> Conserved domain

<400> 216  
Met Ala Ser Ser Ser Ser Ser Tyr Arg Phe Gln Ser Gly Ser Tyr  
1 5 10 15  
Pro Leu Ser Ser Ser Pro Ser Leu Gly Asn Phe Val Glu Arg Ile Lys  
20 25 30  
Asp Ala Cys His Phe Leu Val Ser Ala Val Leu Gly Thr Ile Ile Ser  
35 40 45  
Ala Ile Leu Thr Phe Phe Phe Ala Leu Val Gly Thr Leu Leu Gly Ala  
50 55 60  
Leu Thr Gly Ala Leu Ile Gly Gln Glu Thr Glu Ser Gly Phe Ile Arg  
65 70 75 80  
Gly Ala Ala Ile Gly Ala Ile Ser Gly Ala Val Phe Ser Ile Glu Val  
85 90 95  
Phe Glu Ser Ser Leu Asp Leu Trp Lys Ser Asp Glu Ser Gly Phe Gly  
100 105 110  
Cys Phe Leu Tyr Leu Ile Asp Val Ile Val Ser Leu Leu Ser Gly Arg  
115 120 125  
Leu Val Arg Glu Arg Ile Gly Pro Ala Met Leu Ser Ala Val Gln Ser  
130 135 140  
Gln Met Gly Ala Val Asp Thr Ala Phe Asp Asp His Thr Ser Leu Phe  
145 150 155 160  
Asp Thr Gly Gly Ser Lys Gly Leu Thr Gly Asp Leu Val Glu Lys Ile  
165 170 175  
Pro Lys Met Thr Ile Thr Gly Asn Asn Asn Thr Asp Ala Ser Glu Asn  
180 185 190  
Thr Asp Ser Cys Ser Val Cys Leu Gln Asp Phe Gln Leu Gly Glu Thr  
195 200 205  
Val Arg Ser Leu Pro His Cys His His Met Phe His Leu Pro Cys Ile  
210 215 220  
Asp Asn Trp Leu Leu Arg His Gly Ser Cys Pro Met Cys Arg Arg Asp  
225 230 235 240  
Ile

<210> 217  
<211> 1302  
<212> DNA  
<213> Arabidopsis thaliana

<220>

&lt;221&gt; CDS

&lt;222&gt; (331)...(1149)

&lt;400&gt; 217

tgtctctctc tctggctctc tttctcttaa cgtgatcata acgtgattcg aaaattggat 60  
 atagataggt ttcttggttg atcttgatcc ctctggaaaa ggaggggaga atagcagttc 120  
 atgatgggat ttgtatctg cccgttggag tcacctgcga gattactatg gagtacaagc 180  
 ttcttcgcc ataagatcat gatcttctaa tccttcctac ttcttcccat ctttttaatc 240  
 atcttctcgc tatctctgct tcctctttct ctctgtttcc tctttctcag aactcagaag 300  
 tagttgttgt tttatttctg ttgatcaaaa atg gaa tcc aat tcg ttt ttc ttc 354

Met Glu Ser Asn Ser Phe Phe Phe

1

5

gat cca tct gct tca cac ggc aac agc atg ttc ttc ctt ggg aat ctc 402  
 Asp Pro Ser Ala Ser His Gly Asn Ser Met Phe Phe Leu Gly Asn Leu  
 10 15 20

aat ccc gtc gtc caa gga gga gga gca aga tcg atg atg aac atg gag 450  
 Asn Pro Val Val Gln Gly Gly Gly Ala Arg Ser Met Met Asn Met Glu  
 25 30 35 40

gaa act tcg aag cga agg ccc ttc ttt agc tcc cct gag gat ctc tac 498  
 Glu Thr Ser Lys Arg Arg Pro Phe Phe Ser Ser Pro Glu Asp Leu Tyr  
 45 50 55

gac gat gac ttt tac gac gac cag ttg cct gaa aag aag cgt cgc ctc 546  
 Asp Asp Asp Phe Tyr Asp Asp Gln Leu Pro Glu Lys Lys Arg Arg Leu  
 60 65 70

act acc gaa caa gtg cat ctg ctg gag aaa agc ttc gag aca gag aac 594  
 Thr Thr Glu Gln Val His Leu Leu Glu Lys Ser Phe Glu Thr Glu Asn  
 75 80 85

aag cta gag cct gaa cgc aag act cag ctt gcc aag aag ctt ggt cta 642  
 Lys Leu Glu Pro Glu Arg Lys Thr Gln Leu Ala Lys Lys Leu Gly Leu  
 90 95 100

cag cca agg caa gtg gct gtc tgg ttt cag aat cgc cga gct cgt tgg 690  
 Gln Pro Arg Gln Val Ala Val Trp Phe Gln Asn Arg Arg Ala Arg Trp  
 105 110 115 120

aaa aca aaa cag ctt gag aga gac tac gat ctt ctc aag tcc act tac 738  
 Lys Thr Lys Gln Leu Glu Arg Asp Tyr Asp Leu Leu Lys Ser Thr Tyr  
 125 130 135

gac caa ctt ctt tct aac tac gac tcc atc gtc atg gac aac gat aag 786  
 Asp Gln Leu Leu Ser Asn Tyr Asp Ser Ile Val Met Asp Asn Asp Lys  
 140 145 150

ctc aga tcc gag gtt act tcc ctg acc gaa aag ctt cag ggc aaa caa 834  
 Leu Arg Ser Glu Val Thr Ser Leu Thr Glu Lys Leu Gln Gly Lys Gln  
 155 160 165

gag aca gct aat gaa cca cct ggt caa gtg ccc gaa cca aac caa ctt 882  
 Glu Thr Ala Asn Glu Pro Pro Gly Gln Val Pro Glu Pro Asn Gln Leu  
 170 175 180

```

gat ccg gtt tac att aat gcg gca gca atc aaa acc gag gac cgg tta 930
Asp Pro Val Tyr Ile Asn Ala Ala Ala Ile Lys Thr Glu Asp Arg Leu
185 190 195 200

agt tca ggg agc gtt ggg agc gcg gta cta gac gac gac gca cct caa 978
Ser Ser Gly Ser Val Gly Ser Ala Val Leu Asp Asp Asp Ala Pro Gln
205 210 215

cta cta gac agc tgt gac tct tac ttc cca agc atc gta ccc atc caa 1026
Leu Leu Asp Ser Cys Asp Ser Tyr Phe Pro Ser Ile Val Pro Ile Gln
220 225 230

gac aac agc aac gcc agt gat cat gac aat gac cgg agc tgt ttc gcc 1074
Asp Asn Ser Asn Ala Ser Asp His Asp Asn Asp Arg Ser Cys Phe Ala
235 240 245

gac gtc ttt gtg ccc acc act tca ccg tcg cac gat cat cac ggt gaa 1122
Asp Val Phe Val Pro Thr Thr Ser Pro Ser His Asp His His Gly Glu
250 255 260

tca ttg gct ttc tgg gga tgg cct tag aaaaccactc tgataataaa 1169
Ser Leu Ala Phe Trp Gly Trp Pro *
265 270

tgtgtgttta ttttaagttca agagtcacatc tcttgttgtt tccatgttga cgataattgt 1229
tgactcgtgg aataattccg ctgttcaacg gtatttttat cagttgcatt atatgctttt 1289
atgaaaaaaaa aaa 1302

```

&lt;210&gt; 218

&lt;211&gt; 272

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (64)...(124)

&lt;223&gt; Conserved domain

&lt;400&gt; 218

```

Met Glu Ser Asn Ser Phe Phe Phe Asp Pro Ser Ala Ser His Gly Asn
1 5 10 15
Ser Met Phe Phe Leu Gly Asn Leu Asn Pro Val Val Gln Gly Gly Gly
20 25 30
Ala Arg Ser Met Met Asn Met Glu Glu Thr Ser Lys Arg Arg Pro Phe
35 40 45
Phe Ser Ser Pro Glu Asp Leu Tyr Asp Asp Asp Phe Tyr Asp Asp Gln
50 55 60
Leu Pro Glu Lys Lys Arg Arg Leu Thr Thr Glu Gln Val His Leu Leu
65 70 75 80
Glu Lys Ser Phe Glu Thr Glu Asn Lys Leu Glu Pro Glu Arg Lys Thr
85 90 95
Gln Leu Ala Lys Lys Leu Gly Leu Gln Pro Arg Gln Val Ala Val Trp
100 105 110
Phe Gln Asn Arg Arg Ala Arg Trp Lys Thr Lys Gln Leu Glu Arg Asp
115 120 125
Tyr Asp Leu Leu Lys Ser Thr Tyr Asp Gln Leu Leu Ser Asn Tyr Asp
130 135 140

```

Ser Ile Val Met Asp Asn Asp Lys Leu Arg Ser Glu Val Thr Ser Leu  
 145 150 155 160  
 Thr Glu Lys Leu Gln Gly Lys Gln Glu Thr Ala Asn Glu Pro Pro Gly  
 165 170 175  
 Gln Val Pro Glu Pro Asn Gln Leu Asp Pro Val Tyr Ile Asn Ala Ala  
 180 185 190  
 Ala Ile Lys Thr Glu Asp Arg Leu Ser Ser Gly Ser Val Gly Ser Ala  
 195 200 205  
 Val Leu Asp Asp Asp Ala Pro Gln Leu Leu Asp Ser Cys Asp Ser Tyr  
 210 215 220  
 Phe Pro Ser Ile Val Pro Ile Gln Asp Asn Ser Asn Ala Ser Asp His  
 225 230 235 240  
 Asp Asn Asp Arg Ser Cys Phe Ala Asp Val Phe Val Pro Thr Thr Ser  
 245 250 255  
 Pro Ser His Asp His His Gly Glu Ser Leu Ala Phe Trp Gly Trp Pro  
 260 265 270

<210> 219

<211> 1240

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (88)...(1020)

<400> 219

gtaaatctct ctttgaaggt tcctaactcg ttaatcgtaa ctcacagtga ctcgttcgag 60  
 tcaaagtcctc tgtcttttagc tcaaacc atg gct agt aac aac cct cac gac aac 114  
 Met Ala Ser Asn Asn Pro His Asp Asn  
 1 5

ctt tct gac caa act cct tct gat gat ttc ttc gag caa atc ctc ggc 162  
 Leu Ser Asp Gln Thr Pro Ser Asp Asp Phe Phe Glu Gln Ile Leu Gly  
 10 15 20 25

ctt cct aac ttc tca gcc tct tct gcc gcc ggt tta tct gga gtt gac 210  
 Leu Pro Asn Phe Ser Ala Ser Ser Ala Ala Gly Leu Ser Gly Val Asp  
 30 35 40

gga gga tta ggt ggt gga gca ccg cct atg atg ctg cag ttg ggt tcc 258  
 Gly Gly Leu Gly Gly Gly Ala Pro Pro Met Met Leu Gln Leu Gly Ser  
 45 50 55

gga gaa gaa gga agt cac atg ggt ggc tta gga gga agt gga cca act 306  
 Gly Glu Glu Gly Ser His Met Gly Gly Leu Gly Gly Ser Gly Pro Thr  
 60 65 70

ggg ttt cac aat.cag atg ttt cct ttg ggg tta agt ctt gat caa ggg 354  
 Gly Phe His Asn Gln Met Phe Pro Leu Gly Leu Ser Leu Asp Gln Gly  
 75 80 85

aaa gga cct ggg ttt ctt aga cct gaa gga gga cat gga agt ggg aaa 402  
 Lys Gly Pro Gly Phe Leu Arg Pro Glu Gly Gly His Gly Ser Gly Lys  
 90 95 100 105



```

aga ttc tca gat gat gtt gtt gat aat cga tgt tct tct atg aaa cct 450
Arg Phe Ser Asp Asp Val Val Asp Asn Arg Cys Ser Ser Met Lys Pro
      110      115      120

gtt ttc cac ggg cag cct atg caa cag cca cct cca tgc gcc cca cat 498
Val Phe His Gly Gln Pro Met Gln Gln Pro Pro Pro Ser Ala Pro His
      125      130      135

cag cct act tca atc cgt ccc agg gtt cga gct agg cgt ggt cag gct 546
Gln Pro Thr Ser Ile Arg Pro Arg Val Arg Ala Arg Arg Gly Gln Ala
      140      145      150

act gat cca cat agc atc gct gag cgg cta cgt aga gaa aga ata gca 594
Thr Asp Pro His Ser Ile Ala Glu Arg Leu Arg Arg Glu Arg Ile Ala
      155      160      165

gaa cgg atc agg gcg ctg cag gaa ctt gta cct act gtg aac aag acc 642
Glu Arg Ile Arg Ala Leu Gln Glu Leu Val Pro Thr Val Asn Lys Thr
      170      175      180      185

gat aga gct gct atg atc gat gag att gtc gat tat gta aag ttt ctc 690
Asp Arg Ala Ala Met Ile Asp Glu Ile Val Asp Tyr Val Lys Phe Leu
      190      195      200

agg ctc caa gtc aag gtt ttg agc atg aac cga ctt ggt gga gcc ggt 738
Arg Leu Gln Val Lys Val Leu Ser Met Asn Arg Leu Gly Gly Ala Gly
      205      210      215

gcg gtt gct cca ctt gtt act gat atg cct ctt tca tca tca gtt gag 786
Ala Val Ala Pro Leu Val Thr Asp Met Pro Leu Ser Ser Ser Val Glu
      220      225      230

gat gaa acg ggt gag ggt gga agg act ccg caa cca gcg tgg gag aaa 834
Asp Glu Thr Gly Glu Gly Gly Arg Thr Pro Gln Pro Ala Trp Glu Lys
      235      240      245

tgg tct aac gat ggg act gaa cgt caa gtg gct aaa ctg atg gaa gag 882
Trp Ser Asn Asp Gly Thr Glu Arg Gln Val Ala Lys Leu Met Glu Glu
      250      255      260      265

aac gtt gga gcc gcg atg cag ctt ctt caa tca aag gct ctt tgt atg 930
Asn Val Gly Ala Ala Met Gln Leu Leu Gln Ser Lys Ala Leu Cys Met
      270      275      280

atg cca atc tca ttg gca atg gca att tac cat tct caa cct ccg gat 978
Met Pro Ile Ser Leu Ala Met Ala Ile Tyr His Ser Gln Pro Pro Asp
      285      290      295

aca tct tca gtg gtc aag cct gag aac aat cct cca cag tag 1020
Thr Ser Ser Val Val Lys Pro Glu Asn Asn Pro Pro Gln *
      300      305      310

gatttctgca ataaagagtt tgtacagcta atccaactgt ccaacatggg tttttcttct 1080
gctctaatga ctctggttct ttctctctct tctcaccgac ttgaaaggta aaaaagttaa 1140
aaaggctttg tagatggaat caatgtagga tttgcagtag agggcaaaaa aatgtcatat 1200
agctcaattg atcaagtctt aaaaaaaaaa aaaaaaaaaa 1240

```

<210> 220  
 <211> 310  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <221> DOMAIN  
 <222> (143)... (240)  
 <223> Conserved domain

<400> 220

```

Met Ala Ser Asn Asn Pro His Asp Asn Leu Ser Asp Gln Thr Pro Ser
 1           5           10           15
Asp Asp Phe Phe Glu Gln Ile Leu Gly Leu Pro Asn Phe Ser Ala Ser
      20           25           30
Ser Ala Ala Gly Leu Ser Gly Val Asp Gly Gly Leu Gly Gly Gly Ala
      35           40           45
Pro Pro Met Met Leu Gln Leu Gly Ser Gly Glu Glu Gly Ser His Met
 50           55           60
Gly Gly Leu Gly Gly Ser Gly Pro Thr Gly Phe His Asn Gln Met Phe
65           70           75           80
Pro Leu Gly Leu Ser Leu Asp Gln Gly Lys Gly Pro Gly Phe Leu Arg
      85           90           95
Pro Glu Gly Gly His Gly Ser Gly Lys Arg Phe Ser Asp Asp Val Val
      100          105          110
Asp Asn Arg Cys Ser Ser Met Lys Pro Val Phe His Gly Gln Pro Met
      115          120          125
Gln Gln Pro Pro Pro Ser Ala Pro His Gln Pro Thr Ser Ile Arg Pro
      130          135          140
Arg Val Arg Ala Arg Arg Gly Gln Ala Thr Asp Pro His Ser Ile Ala
145          150          155          160
Glu Arg Leu Arg Arg Glu Arg Ile Ala Glu Arg Ile Arg Ala Leu Gln
      165          170          175
Glu Leu Val Pro Thr Val Asn Lys Thr Asp Arg Ala Ala Met Ile Asp
      180          185          190
Glu Ile Val Asp Tyr Val Lys Phe Leu Arg Leu Gln Val Lys Val Leu
      195          200          205
Ser Met Asn Arg Leu Gly Gly Ala Gly Ala Val Ala Pro Leu Val Thr
210          215          220
Asp Met Pro Leu Ser Ser Ser Val Glu Asp Glu Thr Gly Glu Gly Gly
225          230          235          240
Arg Thr Pro Gln Pro Ala Trp Glu Lys Trp Ser Asn Asp Gly Thr Glu
      245          250          255
Arg Gln Val Ala Lys Leu Met Glu Glu Asn Val Gly Ala Ala Met Gln
      260          265          270
Leu Leu Gln Ser Lys Ala Leu Cys Met Met Pro Ile Ser Leu Ala Met
      275          280          285
Ala Ile Tyr His Ser Gln Pro Pro Asp Thr Ser Ser Val Val Lys Pro
290          295          300
Glu Asn Asn Pro Pro Gln
305          310

```

<210> 221  
 <211> 1431  
 <212> DNA  
 <213> Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (129)... (1211)

&lt;400&gt; 221

```

ctttaaatcc caaaccaacc cttaaagtttt gatttttaaat tttgggggta accaaaaaaaaa 60
aaacaaaacc ctaatTTTTT ttcttttagtg atgagattat tggatgatgat gaaatgattg 120
gagatcta atg aag aat aac aac aat ggc gac gtt gtg gat aac gaa gtg 170
      Met Lys Asn Asn Asn Asn Gly Asp Val Val Asp Asn Glu Val
        1             5             10

aac aac cgg tta agc cgg tgg cat cac aat tct tcc cgg ata att agg 218
Asn Asn Arg Leu Ser Arg Trp His His Asn Ser Ser Arg Ile Ile Arg
  15             20             25             30

gtt tca cga gct tcc ggt ggt aaa gat cga cac agc aaa gtc ttg act 266
Val Ser Arg Ala Ser Gly Gly Lys Asp Arg His Ser Lys Val Leu Thr
          35             40             45

tct aaa gga cca cgt gac cgt cgt gtc cgg tta tca gtc tcc acc gct 314
Ser Lys Gly Pro Arg Asp Arg Arg Val Arg Leu Ser Val Ser Thr Ala
      50             55             60

ctt caa ttc tat gat ctt caa gat cgg tta ggt tat gat caa cct agc 362
Leu Gln Phe Tyr Asp Leu Gln Asp Arg Leu Gly Tyr Asp Gln Pro Ser
      65             70             75

aaa gct gtt gaa tgg tta atc aaa gct gct gaa gat tca atc tct gag 410
Lys Ala Val Glu Trp Leu Ile Lys Ala Ala Glu Asp Ser Ile Ser Glu
      80             85             90

ctt cct tca ctc aac aac act cat ttt ccg acc gat gac gag aat cac 458
Leu Pro Ser Leu Asn Asn Thr His Phe Pro Thr Asp Asp Glu Asn His
      95             100             105             110

cag aat cag aca tta aca aca gtt gct gct aat tcc ttg tct aaa tct 506
Gln Asn Gln Thr Leu Thr Thr Val Ala Ala Asn Ser Leu Ser Lys Ser
          115             120             125

gct tgt agt agc aat tca gac acg agc aag aac tct tct ggt ttg tct 554
Ala Cys Ser Ser Asn Ser Asp Thr Ser Lys Asn Ser Ser Gly Leu Ser
          130             135             140

tta tca aga tcg gag ctt aga gat aaa gct aga gag agg gct aga gag 602
Leu Ser Arg Ser Glu Leu Arg Asp Lys Ala Arg Glu Arg Ala Arg Glu
          145             150             155

aga aca gct aaa gag acc aag gaa aga gat cat aac cac act tcg ttt 650
Arg Thr Ala Lys Glu Thr Lys Glu Arg Asp His Asn His Thr Ser Phe
          160             165             170

acg gat ttg tta aat tcc ggt tca gat ccg gtt aac tca aac cgg caa 698
Thr Asp Leu Leu Asn Ser Gly Ser Asp Pro Val Asn Ser Asn Arg Gln
          175             180             185             190

tgg atg gct tca gct cct tct tca tct cca atg gag tat ttc agt tcg 746

```

```

Trp Met Ala Ser Ala Pro Ser Ser Ser Pro Met Glu Tyr Phe Ser Se
      195                      200                      205

ggg tta att ctc ggg tcg ggt caa caa acc cat ttc cct att tca aca 794
Gly Leu Ile Leu Gly Ser Gly Gln Gln Thr His Phe Pro Ile Ser Thr
      210                      215                      220

aat tct cat cct ttc tca tca atc tcc gat cat cat cat cat cct 842
Asn Ser His Pro Phe Ser Ser Ile Ser Asp His His His His Pro
      225                      230                      235

cat cat cag cat caa gag ttt tca ttc gtt ccc gac cat ttg ata tca 890
His His Gln His Gln Gly Phe Ser Phe Val Pro Asp His Leu Ile Ser
      240                      245                      250

ccg gca gaa tcc aac ggc gga gca ttc aat ctt gat ttt aat atg tca 938
Pro Ala Glu Ser Asn Gly Gly Ala Phe Asn Leu Asp Phe Asn Met Ser
      255                      260                      265                      270

aca ccc tcc ggc gcc gga gct gcc gtc tcc gcc gca tca ggt ggt ggc 986
Thr Pro Ser Gly Ala Gly Ala Ala Val Ser Ala Ala Ser Gly Gly Gly
      275                      280                      285

ttc agt ggt ttc aac agg ggg acc ctt cag tcc aat tca aca aat cag 1034
Phe Ser Gly Phe Asn Arg Gly Thr Leu Gln Ser Asn Ser Thr Asn Gln
      290                      295                      300

cat cag tca ttc ctc gct aat cta cag agg ttt cca aca tca gaa agt 1082
His Gln Ser Phe Leu Ala Asn Leu Gln Arg Phe Pro Thr Ser Glu Ser
      305                      310                      315

gga gga ggt cca cag ttc tta ttc ggt gca ctg cct gca gag aat cac 1130
Gly Gly Gly Pro Gln Phe Leu Phe Gly Ala Leu Pro Ala Glu Asn His
      320                      325                      330

cac cac aat cac cag ttt cag ctt tac tat gaa aat gga tgc aga aac 1178
His His Asn His Gln Phe Gln Leu Tyr Tyr Glu Asn Gly Cys Arg Asn
      335                      340                      345                      350

tca tca gaa cat aag ggt aaa ggc aag aac tga tgatattaat tattgcatct 1231
Ser Ser Glu His Lys Gly Lys Gly Lys Asn *
      355                      360

ttggttttgt tcaaatgctc attttgtatg tttatctttg gtttatttca aaacaaatgt 1291
taatctcttt cgttgtctga tgtgtgtag ggttttgttt tatgtattga gggctctttgg 1351
aaatcttttt gcattgtgct tgtaatgttg tatttgtgat aatagcattt tgtttgtgag 1411
ttaaaaaaaaa aaaaaaaaaa 1431

<210> 222
<211> 360
<212> PRT
<213> Arabidopsis thaliana

<220>
<221> DOMAIN
<222> (39)...(95)
<223> Conserved domain

```

&lt;400&gt; 222

```

Met Lys Asn Asn Asn Asn Gly Asp Val Val Asp Asn Glu Val Asn Asn
 1           5           10           15
Arg Leu Ser Arg Trp His His Asn Ser Ser Arg Ile Ile Arg Val Ser
          20           25           30
Arg Ala Ser Gly Gly Lys Asp Arg His Ser Lys Val Leu Thr Ser Lys
          35           40           45
Gly Pro Arg Asp Arg Arg Val Arg Leu Ser Val Ser Thr Ala Leu Gln
          50           55           60
Phe Tyr Asp Leu Gln Asp Arg Leu Gly Tyr Asp Gln Pro Ser Lys Ala
65           70           75           80
Val Glu Trp Leu Ile Lys Ala Ala Glu Asp Ser Ile Ser Glu Leu Pro
          85           90           95
Ser Leu Asn Asn Thr His Phe Pro Thr Asp Asp Glu Asn His Gln Asn
          100          105          110
Gln Thr Leu Thr Thr Val Ala Ala Asn Ser Leu Ser Lys Ser Ala Cys
          115          120          125
Ser Ser Asn Ser Asp Thr Ser Lys Asn Ser Ser Gly Leu Ser Leu Ser
          130          135          140
Arg Ser Glu Leu Arg Asp Lys Ala Arg Glu Arg Ala Arg Glu Arg Thr
145          150          155          160
Ala Lys Glu Thr Lys Glu Arg Asp His Asn His Thr Ser Phe Thr Asp
          165          170          175
Leu Leu Asn Ser Gly Ser Asp Pro Val Asn Ser Asn Arg Gln Trp Met
          180          185          190
Ala Ser Ala Pro Ser Ser Ser Pro Met Glu Tyr Phe Ser Ser Gly Leu
          195          200          205
Ile Leu Gly Ser Gly Gln Gln Thr His Phe Pro Ile Ser Thr Asn Ser
          210          215          220
His Pro Phe Ser Ser Ile Ser Asp His His His His His Pro His His
225          230          235          240
Gln His Gln Glu Phe Ser Phe Val Pro Asp His Leu Ile Ser Pro Ala
          245          250          255
Glu Ser Asn Gly Gly Ala Phe Asn Leu Asp Phe Asn Met Ser Thr Pro
          260          265          270
Ser Gly Ala Gly Ala Ala Val Ser Ala Ala Ser Gly Gly Gly Phe Ser
          275          280          285
Gly Phe Asn Arg Gly Thr Leu Gln Ser Asn Ser Thr Asn Gln His Gln
          290          295          300
Ser Phe Leu Ala Asn Leu Gln Arg Phe Pro Thr Ser Glu Ser Gly Gly
305          310          315          320
Gly Pro Gln Phe Leu Phe Gly Ala Leu Pro Ala Glu Asn His His His
          325          330          335
Asn His Gln Phe Gln Leu Tyr Tyr Glu Asn Gly Cys Arg Asn Ser Ser
          340          345          350
Glu His Lys Gly Lys Gly Lys Asn
          355          360

```

&lt;210&gt; 223

&lt;211&gt; 1571

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (428)...(1402)

&lt;400&gt; 223

```

aggaacagtg aaagggttcgg ttttttgggt ttcgatctga taatcaacaa gaaaaaaggg 60
tttgatttat gtcggctggg tttgaatcga ctgtgatttt gtctttgatt catatctctt 120
ctccgatttc atcatcatct tccccatcat cgtcgtcttt gaaatcttgt cttctcaacg 180
ctcttcactt ctgctgtaat aagcagaggc ttgttctgga gactccttct ctttccatgc 240
gcttaagacc caaaaggact tgttctagt ttgaagtctt tgggggtttt cacataaage 300
agcaaaagtt ttcttttttc atagttcgct gagagttttg agttttgata ccaaaaaagt 360
tttgaccttt tagagtgatt ttttgttctt tctgttttct ggggtattttt gaggagtggg 420
ttaaaca atg gtt gcg att aga aag gaa cag tct ttg agt ggt gtt agt 469
      Met Val Ala Ile Arg Lys Glu Gln Ser Leu Ser Gly Val Ser
        1             5             10

agc gag att aag aag aga gct aag aga aac act cta tcg tcc ctt cct 517
Ser Glu Ile Lys Lys Arg Ala Lys Arg Asn Thr Leu Ser Ser Leu Pro
  15             20             25             30

caa gaa acc caa cct ttg agg aaa gtc cgt att att gtg aat gat cct 565
Gln Glu Thr Gln Pro Leu Arg Lys Val Arg Ile Ile Val Asn Asp Pro
          35             40             45

tat gct act gat gat tcc tct agt gat gag gaa gag ctt aag gtt cct 613
Tyr Ala Thr Asp Asp Ser Ser Ser Asp Glu Glu Glu Leu Lys Val Pro
          50             55             60

aag cca agg aaa atg aaa cgt atc gtt cgt gag att aac ttt cct tct 661
Lys Pro Arg Lys Met Lys Arg Ile Val Arg Glu Ile Asn Phe Pro Ser
          65             70             75

atg gaa gtt tct gaa cag cct tct gag agt tct tct cag gac agt act 709
Met Glu Val Ser Glu Gln Pro Ser Glu Ser Ser Ser Gln Asp Ser Thr
          80             85             90

aaa act gat ggc aag ata gct gtg tca gct tct cct gct gtt cct agg 757
Lys Thr Asp Gly Lys Ile Ala Val Ser Ala Ser Pro Ala Val Pro Arg
          95             100             105             110

aag aag cct gtt ggt gtt agg caa agg aaa tgg ggg aaa tgg gct gct 805
Lys Lys Pro Val Gly Val Arg Gln Arg Lys Trp Gly Lys Trp Ala Ala
          115             120             125

gag att aga gat cct att aag aaa act agg act tgg ttg ggt act ttt 853
Glu Ile Arg Asp Pro Ile Lys Lys Thr Arg Thr Trp Leu Gly Thr Phe
          130             135             140

gat act ctt gaa gaa gct gct aaa gct tat gat gct aag aag ctt gag 901
Asp Thr Leu Glu Glu Ala Ala Lys Ala Tyr Asp Ala Lys Lys Leu Glu
          145             150             155

ttt gat gct att gtt gct gga aat gtg tcc act act aaa cgt gat gtt 949
Phe Asp Ala Ile Val Ala Gly Asn Val Ser Thr Thr Lys Arg Asp Val
          160             165             170

tct tca tct gag act agc caa tgc tct cgt tct tca cct gtt gtt cct 997
Ser Ser Ser Glu Thr Ser Gln Cys Ser Arg Ser Ser Pro Val Val Pro
          175             180             185             190

```

```

gtt gag caa gat gac act tct gca tca gct ctc act tgt gtc aac aac 1045
Val Glu Gln Asp Asp Thr Ser Ala Ser Ala Leu Thr Cys Val Asn Asn
      195                200                205

cct gat gac gtc tcg acc gtt gct cca act gct cca act cca aat gtt 1093
Pro Asp Asp Val Ser Thr Val Ala Pro Thr Ala Pro Thr Pro Asn Val
      210                215                220

cct gct ggt gga aac aag gaa acg ttg ttc gat ttc gac ttt act aat 1141
Pro Ala Gly Gly Asn Lys Glu Thr Leu Phe Asp Phe Asp Phe Thr Asn
      225                230                235

cta cag atc cct gat ttt ggt ttc ttg gca gag gag caa caa gac cta 1189
Leu Gln Ile Pro Asp Phe Gly Phe Leu Ala Glu Glu Gln Gln Asp Leu
      240                245                250

gac ttc gat tgt ttc ctc gcg gat gat cag ttt gat gat ttc ggc ttg 1237
Asp Phe Asp Cys Phe Leu Ala Asp Asp Gln Phe Asp Asp Phe Gly Leu
      255                260                265                270

ctt gat gac att caa gga ttc gaa gat aac ggt cca agt gcg tta cca 1285
Leu Asp Asp Ile Gln Gly Phe Glu Asp Asn Gly Pro Ser Ala Leu Pro
      275                280                285

gat ttc gac ttt gcg gat gtt gaa gat ctt cag cta gct gac tct agt 1333
Asp Phe Asp Phe Ala Asp Val Glu Asp Leu Gln Leu Ala Asp Ser Ser
      290                295                300

ttc ggt ttc ctt gat caa ctt gct cct atc aac atc tct tgc cca tta 1381
Phe Gly Phe Leu Asp Gln Leu Ala Pro Ile Asn Ile Ser Cys Pro Leu
      305                310                315

aaa agt ttt gca gct tca tag gatcttgctt agtaatgtta agtgagaaga 1432
Lys Ser Phe Ala Ala Ser *
      320

gtgttttggtt ttttcgttta tgcttttagta atttaagaca tacaaaagtg tgtgttccgg 1492
attgtagtaa gatcttaaga cataaagccg gggtttgcaa ttaggaatcg agttttaatg 1552
aagtttttagt ttatgtttg 1571

```

&lt;210&gt; 224

&lt;211&gt; 324

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (109)...(177)

&lt;223&gt; Conserved domain

&lt;400&gt; 224

```

Met Val Ala Ile Arg Lys Glu Gln Ser Leu Ser Gly Val Ser Ser Glu
 1           5           10           15
Ile Lys Lys Arg Ala Lys Arg Asn Thr Leu Ser Ser Leu Pro Gln Glu
      20           25           30
Thr Gln Pro Leu Arg Lys Val Arg Ile Ile Val Asn Asp Pro Tyr Ala

```

```

      35              40              45
Thr Asp Asp Ser Ser Ser Asp Glu Glu Glu Leu Lys Val Pro Lys Pro
  50              55              60
Arg Lys Met Lys Arg Ile Val Arg Glu Ile Asn Phe Pro Ser Met Glu
  65              70              75              80
Val Ser Glu Gln Pro Ser Glu Ser Ser Ser Gln Asp Ser Thr Lys Thr
      85              90              95
Asp Gly Lys Ile Ala Val Ser Ala Ser Pro Ala Val Pro Arg Lys Lys
      100              105              110
Pro Val Gly Val Arg Gln Arg Lys Trp Gly Lys Trp Ala Ala Glu Ile
      115              120              125
Arg Asp Pro Ile Lys Lys Thr Arg Thr Trp Leu Gly Thr Phe Asp Thr
      130              135              140
Leu Glu Glu Ala Ala Lys Ala Tyr Asp Ala Lys Lys Leu Glu Phe Asp
      145              150              155              160
Ala Ile Val Ala Gly Asn Val Ser Thr Thr Lys Arg Asp Val Ser Ser
      165              170              175
Ser Glu Thr Ser Gln Cys Ser Arg Ser Ser Pro Val Val Pro Val Glu
      180              185              190
Gln Asp Asp Thr Ser Ala Ser Ala Leu Thr Cys Val Asn Asn Pro Asp
      195              200              205
Asp Val Ser Thr Val Ala Pro Thr Ala Pro Thr Pro Asn Val Pro Ala
      210              215              220
Gly Gly Asn Lys Glu Thr Leu Phe Asp Phe Asp Phe Thr Asn Leu Gln
      225              230              235              240
Ile Pro Asp Phe Gly Phe Leu Ala Glu Glu Gln Gln Asp Leu Asp Phe
      245              250              255
Asp Cys Phe Leu Ala Asp Asp Gln Phe Asp Asp Phe Gly Leu Leu Asp
      260              265              270
Asp Ile Gln Gly Phe Glu Asp Asn Gly Pro Ser Ala Leu Pro Asp Phe
      275              280              285
Asp Phe Ala Asp Val Glu Asp Leu Gln Leu Ala Asp Ser Ser Phe Gly
      290              295              300
Phe Leu Asp Gln Leu Ala Pro Ile Asn Ile Ser Cys Pro Leu Lys Ser
      305              310              315              320
Phe Ala Ala Ser

```

&lt;210&gt; 225

&lt;211&gt; 1152

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (76)...(1008)

&lt;400&gt; 225

```

gtgaccgaag aaagcaaatt gagactacgc accaactagt cctttgggttt gtatcttaag 60
ataaagggttt cttttt atg gac ggt tct tcg ttt ctc gac atc tct ctc gat 111
      Met Asp Gly Ser Ser Phe Leu Asp Ile Ser Leu Asp
              1              5              10

```

```

ctc aac acc aat cct ttc tcc gca aaa ctt ccg aag aag gag gtc tca 159
Leu Asn Thr Asn Pro Phe Ser Ala Lys Leu Pro Lys Lys Glu Val Ser
      15              20              25

```



gtt ttg gct tct act cac tta aag agg aaa tgg ttg gag caa gac gag	207
Val Leu Ala Ser Thr His Leu Lys Arg Lys Trp Leu Glu Gln Asp Glu	
30 35 40	
agc gca agt gag tta cga gag gag cta aac aga gtt aat tca gag aac	255
Ser Ala Ser Glu Leu Arg Glu Glu Leu Asn Arg Val Asn Ser Glu Asn	
45 50 55 60	
aag aag cta aca gag atg tta gct aga gtc tgt gag agc tac aac gaa	303
Lys Lys Leu Thr Glu Met Leu Ala Arg Val Cys Glu Ser Tyr Asn Glu	
65 70 75	
cta cat aat cat ttg gag aag ctt cag agt cgc cag agc cct gaa atc	351
Leu His Asn His Leu Glu Lys Leu Gln Ser Arg Gln Ser Pro Glu Ile	
80 85 90	
gag cag acc gat ata ccg ata aag aaa aga aaa caa gac ccg gat gag	399
Glu Gln Thr Asp Ile Pro Ile Lys Lys Arg Lys Gln Asp Pro Asp Glu	
95 100 105	
ttc tta ggc ttt cct att gga ctc agt agt gga aaa act gag aac agc	447
Phe Leu Gly Phe Pro Ile Gly Leu Ser Ser Gly Lys Thr Glu Asn Ser	
110 115 120	
tcc agc aac gaa gat cat cat cat cat cat cag caa cat gag cag aaa	495
Ser Ser Asn Glu Asp His His His His His Gln Gln His Glu Gln Lys	
125 130 135 140	
aat cag ctt ctt tca tgt aaa aga cca gtc act gat agc ttc aac aaa	543
Asn Gln Leu Leu Ser Cys Lys Arg Pro Val Thr Asp Ser Phe Asn Lys	
145 150 155	
gca aaa gtt tcg act gtc tac gtg cct act gaa aca tcg gac aca agc	591
Ala Lys Val Ser Thr Val Tyr Val Pro Thr Glu Thr Ser Asp Thr Ser	
160 165 170	
ttg aca gtt aaa gat gga ttt caa tgg agg aaa tac gga caa aag gtt	639
Leu Thr Val Lys Asp Gly Phe Gln Trp Arg Lys Tyr Gly Gln Lys Val	
175 180 185	
aca aga gac aac ccg tca cct aga gct tac ttt aga tgc tcg ttt gca	687
Thr Arg Asp Asn Pro Ser Pro Arg Ala Tyr Phe Arg Cys Ser Phe Ala	
190 195 200	
ccg tct tgt cca gta aaa aag aag gta caa cgc agc gca gag gat cca	735
Pro Ser Cys Pro Val Lys Lys Lys Val Gln Arg Ser Ala Glu Asp Pro	
205 210 215 220	
tct tta ctt gta gcg aca tac gaa ggg acg cat aac cac ttg ggt cca	783
Ser Leu Leu Val Ala Thr Tyr Glu Gly Thr His Asn His Leu Gly Pro	
225 230 235	
aat gct tct gaa ggg gat gct aca agc cag ggt ggg tca agc aca gtg	831
Asn Ala Ser Glu Gly Asp Ala Thr Ser Gln Gly Gly Ser Ser Thr Val	
240 245 250	

act ttg gat ctg gtt aat ggc tgt cat aga cta gcg ttg gag aaa aa  
 Thr Leu Asp Leu Val Asn Gly Cys His Arg Leu Ala Leu Glu Lys Asn  
 255 260 265

gaa agg gat aat acg atg caa gag gtt ctg att caa caa atg gcg tca 927  
 Glu Arg Asp Asn Thr Met Gln Glu Val Leu Ile Gln Gln Met Ala Ser  
 270 275 280

tcg tta aca aaa gat tcg aaa ttt aca gct gct ctt gct gct gct ata 975  
 Ser Leu Thr Lys Asp Ser Lys Phe Thr Ala Ala Leu Ala Ala Ala Ile  
 285 290 295 300

tct ggg agg tta atg gag caa tct aga aca tga acgttttttag tgaatgtatt 1028  
 Ser Gly Arg Leu Met Glu Gln Ser Arg Thr \*  
 305 310

gtttttgttt gtttagaatg attcttcggtt ttcgaattgt gtcttttcgat taggagataa 1088  
 aagatgtata taaatattat aagtagatga agaaatcgta taagtaaaaa aaaaaaaaaa 1148  
 aaaa 1152

&lt;210&gt; 226

&lt;211&gt; 310

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (176)...(233)

&lt;223&gt; Conserved domain

&lt;400&gt; 226

Met Asp Gly Ser Ser Phe Leu Asp Ile Ser Leu Asp Leu Asn Thr Asn  
 1 5 10 15  
 Pro Phe Ser Ala Lys Leu Pro Lys Lys Glu Val Ser Val Leu Ala Ser  
 20 25 30  
 Thr His Leu Lys Arg Lys Trp Leu Glu Gln Asp Glu Ser Ala Ser Glu  
 35 40 45  
 Leu Arg Glu Glu Leu Asn Arg Val Asn Ser Glu Asn Lys Lys Leu Thr  
 50 55 60  
 Glu Met Leu Ala Arg Val Cys Glu Ser Tyr Asn Glu Leu His Asn His  
 65 70 75 80  
 Leu Glu Lys Leu Gln Ser Arg Gln Ser Pro Glu Ile Glu Gln Thr Asp  
 85 90 95  
 Ile Pro Ile Lys Lys Arg Lys Gln Asp Pro Asp Glu Phe Leu Gly Phe  
 100 105 110  
 Pro Ile Gly Leu Ser Ser Gly Lys Thr Glu Asn Ser Ser Ser Asn Glu  
 115 120 125  
 Asp His His His His Gln Gln His Glu Gln Lys Asn Gln Leu Leu  
 130 135 140  
 Ser Cys Lys Arg Pro Val Thr Asp Ser Phe Asn Lys Ala Lys Val Ser  
 145 150 155 160  
 Thr Val Tyr Val Pro Thr Glu Thr Ser Asp Thr Ser Leu Thr Val Lys  
 165 170 175  
 Asp Gly Phe Gln Trp Arg Lys Tyr Gly Gln Lys Val Thr Arg Asp Asn  
 180 185 190  
 Pro Ser Pro Arg Ala Tyr Phe Arg Cys Ser Phe Ala Pro Ser Cys Pro  
 195 200 205

Val Lys Lys Lys Val Gln Arg Ser Ala Glu Asp Pro Ser Leu Leu Va  
 210 215 220  
 Ala Thr Tyr Glu Gly Thr His Asn His Leu Gly Pro Asn Ala Ser Glu  
 225 230 235 240  
 Gly Asp Ala Thr Ser Gln Gly Gly Ser Ser Thr Val Thr Leu Asp Leu  
 245 250 255  
 Val Asn Gly Cys His Arg Leu Ala Leu Glu Lys Asn Glu Arg Asp Asn  
 260 265 270  
 Thr Met Gln Glu Val Leu Ile Gln Gln Met Ala Ser Ser Leu Thr Lys  
 275 280 285  
 Asp Ser Lys Phe Thr Ala Ala Leu Ala Ala Ile Ser Gly Arg Leu  
 290 295 300  
 Met Glu Gln Ser Arg Thr  
 305 310

&lt;210&gt; 227

&lt;211&gt; 715

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (20)...(703)

&lt;400&gt; 227

caatttttaa ccggtcatt atg atg tct act gta ccg gcc ttt acc ttt act 52  
 Met Met Ser Thr Val Pro Ala Phe Thr Phe Thr  
 1 5 10

gaa ccg ggt ttg gtt aac caa tta tcg gat ttc cag acc gga ttc act 100  
 Glu Pro Gly Leu Val Asn Gln Leu Ser Asp Phe Gln Thr Gly Phe Thr  
 15 20 25

cct tgg gaa ttg aac tgc tcc gat ctc ttc tct aca atc cat ctc gaa 148  
 Pro Trp Glu Leu Asn Cys Ser Asp Leu Phe Ser Thr Ile His Leu Glu  
 30 35 40

ccg gtc gta ccg agt cct tgt tct ggt gaa tcc gat gcc ggt tct gtc 196  
 Pro Val Val Pro Ser Pro Cys Ser Gly Glu Ser Asp Ala Gly Ser Val  
 45 50 55

aaa att aac acc gat ttt aac ggt ttt gac gaa tcg tgt atc ggt tcc 244  
 Lys Ile Asn Thr Asp Phe Asn Gly Phe Asp Glu Ser Cys Ile Gly Ser  
 60 65 70 75

atc aaa act aac tcc ggt tct gat gat tcc aac ctt ttc cac ggc gta 292  
 Ile Lys Thr Asn Ser Gly Ser Asp Asp Ser Asn Leu Phe His Gly Val  
 80 85 90

ccg agt cct caa tcc gac gaa ttg gac tca aaa aac acg aaa atc cga 340  
 Pro Ser Pro Gln Ser Asp Glu Leu Asp Ser Lys Asn Thr Lys Ile Arg  
 95 100 105

agt aac gcc acg aat cat aac cgg aac aaa ttg aac cgg tcg gtt ttg 388  
 Ser Asn Ala Thr Asn His Asn Arg Asn Lys Leu Asn Arg Ser Val Leu  
 110 115 120

```

cag gtg act gac gac cgt aaa cgc aaa cgg atg gaa tca aac cga gaa 436
Gln Val Thr Asp Asp Arg Lys Arg Lys Arg Met Glu Ser Asn Arg Glu
    125                130                135

tca gcg aag cgg tcg agg atg cgt aaa caa aga cac att gat aat tta 484
Ser Ala Lys Arg Ser Arg Met Arg Lys Gln Arg His Ile Asp Asn Leu
    140                145                150                155

aaa gac gaa gca aat cgt ctc ggt tta gaa aac cgg gaa ctc gca aac 532
Lys Asp Glu Ala Asn Arg Leu Gly Leu Glu Asn Arg Glu Leu Ala Asn
                160                165                170

cgg ctt cga att gtt ttg tac aac atc gca tta atg tgt acg gac aac 580
Arg Leu Arg Ile Val Leu Tyr Asn Ile Ala Leu Met Cys Thr Asp Asn
                175                180                185

aat cag ctt ttg tcg gaa caa gag att ctc aga cgg aga ttc ttg gag 628
Asn Gln Leu Leu Ser Glu Gln Glu Ile Leu Arg Arg Arg Phe Leu Glu
    190                195                200

atg agg cag att ttg att ttc aga cag ctt cag ctg aat cca tca ttg 676
Met Arg Gln Ile Leu Ile Phe Arg Gln Leu Gln Leu Asn Pro Ser Leu
    205                210                215

atc atc aat cat cat cat atg att tga aagaaaaaaaa aa 715
Ile Ile Asn His His His Met Ile *
    220                225

```

&lt;210&gt; 228

&lt;211&gt; 227

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (129)...(180)

&lt;223&gt; Conserved domain

&lt;400&gt; 228

```

Met Met Ser Thr Val Pro Ala Phe Thr Phe Thr Glu Pro Gly Leu Val
  1                5                10                15
Asn Gln Leu Ser Asp Phe Gln Thr Gly Phe Thr Pro Trp Glu Leu Asn
    20                25                30
Cys Ser Asp Leu Phe Ser Thr Ile His Leu Glu Pro Val Val Pro Ser
    35                40                45
Pro Cys Ser Gly Glu Ser Asp Ala Gly Ser Val Lys Ile Asn Thr Asp
    50                55                60
Phe Asn Gly Phe Asp Glu Ser Cys Ile Gly Ser Ile Lys Thr Asn Ser
    65                70                75                80
Gly Ser Asp Asp Ser Asn Leu Phe His Gly Val Pro Ser Pro Gln Ser
    85                90                95
Asp Glu Leu Asp Ser Lys Asn Thr Lys Ile Arg Ser Asn Ala Thr Asn
    100                105                110
His Asn Arg Asn Lys Leu Asn Arg Ser Val Leu Gln Val Thr Asp Asp
    115                120                125

```

Arg Lys Arg Lys Arg Met Glu Ser Asn Arg Glu Ser Ala Lys Arg Se  
 130 135 140  
 Arg Met Arg Lys Gln Arg His Ile Asp Asn Leu Lys Asp Glu Ala Asn  
 145 150 155 160  
 Arg Leu Gly Leu Glu Asn Arg Glu Leu Ala Asn Arg Leu Arg Ile Val  
 165 170 175  
 Leu Tyr Asn Ile Ala Leu Met Cys Thr Asp Asn Asn Gln Leu Leu Ser  
 180 185 190  
 Glu Gln Glu Ile Leu Arg Arg Arg Phe Leu Glu Met Arg Gln Ile Leu  
 195 200 205  
 Ile Phe Arg Gln Leu Gln Leu Asn Pro Ser Leu Ile Ile Asn His His  
 210 215 220  
 His Met Ile  
 225

<210> 229

<211> 1011

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (57)...(983)

<400> 229

cgtctaccta ctatggtctg gagattagtt cgtttattga actaatgttt tagaca atg 59  
 Met  
 1

caa gag ttc cat agt agc aaa gat tca ttg cct tgt cct gca act tct 107  
 Gln Glu Phe His Ser Ser Lys Asp Ser Leu Pro Cys Pro Ala Thr Ser  
 5 10 15

tgg gat aac tct gtc ttc acc aac tca aat gtc caa gga tca tca tcc 155  
 Trp Asp Asn Ser Val Phe Thr Asn Ser Asn Val Gln Gly Ser Ser Ser  
 20 25 30

ttg acc gat aac aac act tta agc ttg aca atg gag atg aaa caa act 203  
 Leu Thr Asp Asn Asn Thr Leu Ser Leu Thr Met Glu Met Lys Gln Thr  
 35 40 45

ggt ttt caa atg cag cac tat gat tcc tcc tct act caa tcc act gga 251  
 Gly Phe Gln Met Gln His Tyr Asp Ser Ser Ser Thr Gln Ser Thr Gly  
 50 55 60 65

gga gaa tca tat agt gaa gtt gct agc tta agt gaa cct act aat cgt 299  
 Gly Glu Ser Tyr Ser Glu Val Ala Ser Leu Ser Glu Pro Thr Asn Arg  
 70 75 80

tat ggc cac aac att gtt gtc act cat ctc tca ggt tac aaa gaa aac 347  
 Tyr Gly His Asn Ile Val Val Thr His Leu Ser Gly Tyr Lys Glu Asn  
 85 90 95

ccg gaa aat cct att gga agt cat tcg ata tca aag gtg tct caa gat 395  
 Pro Glu Asn Pro Ile Gly Ser His Ser Ile Ser Lys Val Ser Gln Asp  
 100 105 110

tca gtg gtt ctt cct att gag gcg gct tct tgg cct tta cac ggc aat 443  
 Ser Val Val Leu Pro Ile Glu Ala Ala Ser Trp Pro Leu His Gly Asn  
 115 120 125

gta acg cca cat ttc aat ggt ttc ttg tct ttt cct tat gca tca caa 491  
 Val Thr Pro His Phe Asn Gly Phe Leu Ser Phe Pro Tyr Ala Ser Gln  
 130 135 140 145

cac acg gtg cag cat cct caa atc aga ggg ttg gtt ccg tct aga atg 539  
 His Thr Val Gln His Pro Gln Ile Arg Gly Leu Val Pro Ser Arg Met  
 150 155 160

cct ttg cct cac aac att cca gag aac gaa cca att ttc gtc aat gca 587  
 Pro Leu Pro His Asn Ile Pro Glu Asn Glu Pro Ile Phe Val Asn Ala  
 165 170 175

aaa cag tac caa gcc att ctc cgc cgc aga gag cgc cgt gca aag ctt 635  
 Lys Gln Tyr Gln Ala Ile Leu Arg Arg Arg Glu Arg Ala Lys Leu  
 180 185 190

gaa gct cag aac aag ctc atc aaa gtc cgc aaa cca tat ctt cac gag 683  
 Glu Ala Gln Asn Lys Leu Ile Lys Val Arg Lys Pro Tyr Leu His Glu  
 195 200 205

tcg cgg cac ctc cat gca cta aag aga gtt aga ggc tct ggt gga cgt 731  
 Ser Arg His Leu His Ala Leu Lys Arg Val Arg Gly Ser Gly Gly Arg  
 210 215 220 225

ttc ctc aac aca aag aag cat caa gaa tca aat tcc tca cta tct cct 779  
 Phe Leu Asn Thr Lys Lys His Gln Glu Ser Asn Ser Ser Leu Ser Pro  
 230 235 240

cca ttc ttg att cca cct cat gtc ttc aag aac tct cca gga aag ttc 827  
 Pro Phe Leu Ile Pro Pro His Val Phe Lys Asn Ser Pro Gly Lys Phe  
 245 250 255

cgg caa atg gac att tca agg ggt ggg gtt gtg tct agt gtc tcg aca 875  
 Arg Gln Met Asp Ile Ser Arg Gly Gly Val Val Ser Ser Val Ser Thr  
 260 265 270

aca tct tgc tcg gac ata acc ggg aac aac aac gac atg ttc cag caa 923  
 Thr Ser Cys Ser Asp Ile Thr Gly Asn Asn Asn Asp Met Phe Gln Gln  
 275 280 285

aac cca caa ttc agg ttc tca ggt tat cca tca aac cac cat gtc tca 971  
 Asn Pro Gln Phe Arg Phe Ser Gly Tyr Pro Ser Asn His His Val Ser  
 290 295 300 305

gtc ctc atg tga gagagctccc gcaagtgggtg gatgagggc 1011  
 Val Leu Met \*

&lt;210&gt; 230

&lt;211&gt; 308

&lt;212&gt; PRT

<213> Arabidopsis thaliana

<220>

<221> DOMAIN

<222> (174)...(226)

<223> Conserved domain

<400> 230

```

Met Gln Glu Phe His Ser Ser Lys Asp Ser Leu Pro Cys Pro Ala Thr
 1          5          10          15
Ser Trp Asp Asn Ser Val Phe Thr Asn Ser Asn Val Gln Gly Ser Ser
 20          25          30
Ser Leu Thr Asp Asn Asn Thr Leu Ser Leu Thr Met Glu Met Lys Gln
 35          40          45
Thr Gly Phe Gln Met Gln His Tyr Asp Ser Ser Ser Thr Gln Ser Thr
 50          55          60
Gly Gly Glu Ser Tyr Ser Glu Val Ala Ser Leu Ser Glu Pro Thr Asn
 65          70          75          80
Arg Tyr Gly His Asn Ile Val Val Thr His Leu Ser Gly Tyr Lys Glu
 85          90          95
Asn Pro Glu Asn Pro Ile Gly Ser His Ser Ile Ser Lys Val Ser Gln
100          105          110
Asp Ser Val Val Leu Pro Ile Glu Ala Ala Ser Trp Pro Leu His Gly
115          120          125
Asn Val Thr Pro His Phe Asn Gly Phe Leu Ser Phe Pro Tyr Ala Ser
130          135          140
Gln His Thr Val Gln His Pro Gln Ile Arg Gly Leu Val Pro Ser Arg
145          150          155          160
Met Pro Leu Pro His Asn Ile Pro Glu Asn Glu Pro Ile Phe Val Asn
165          170          175
Ala Lys Gln Tyr Gln Ala Ile Leu Arg Arg Arg Glu Arg Arg Ala Lys
180          185          190
Leu Glu Ala Gln Asn Lys Leu Ile Lys Val Arg Lys Pro Tyr Leu His
195          200          205
Glu Ser Arg His Leu His Ala Leu Lys Arg Val Arg Gly Ser Gly Gly
210          215          220
Arg Phe Leu Asn Thr Lys Lys His Gln Glu Ser Asn Ser Ser Leu Ser
225          230          235          240
Pro Pro Phe Leu Ile Pro Pro His Val Phe Lys Asn Ser Pro Gly Lys
245          250          255
Phe Arg Gln Met Asp Ile Ser Arg Gly Gly Val Val Ser Ser Val Ser
260          265          270
Thr Thr Ser Cys Ser Asp Ile Thr Gly Asn Asn Asn Asp Met Phe Gln
275          280          285
Gln Asn Pro Gln Phe Arg Phe Ser Gly Tyr Pro Ser Asn His His Val
290          295          300
Ser Val Leu Met
305

```

<210> 231

<211> 914

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

&lt;222&gt; (37)...(672)

&lt;400&gt; 231

```

aaatcatcag atagaaggaa atattctgat tgagag atg gct cgt gga aag att 54
                               Met Ala Arg Gly Lys Ile
                               1       5

cag ctt aag agg att gag aac ccg gtt cac aga caa gtg act ttt tgc 102
Gln Leu Lys Arg Ile Glu Asn Pro Val His Arg Gln Val Thr Phe Cys
      10              15              20

aag agg aga act ggt ctt ctc aag aag gct aag gag ctc tct gtg ctc 150
Lys Arg Arg Thr Gly Leu Leu Lys Lys Ala Lys Glu Leu Ser Val Leu
      25              30              35

tgt gat gcc gag atc ggt gtt gtg atc ttc tct cct cag ggc aag ctc 198
Cys Asp Ala Glu Ile Gly Val Val Ile Phe Ser Pro Gln Gly Lys Leu
      40              45              50

ttt gag ctc gct act aaa gga aca atg gag gga atg att gat aag tac 246
Phe Glu Leu Ala Thr Lys Gly Thr Met Glu Gly Met Ile Asp Lys Tyr
      55              60              65              70

atg aag tgt act ggt ggt ggt cgt ggt tct tct tct gct act ttt act 294
Met Lys Cys Thr Gly Gly Gly Arg Gly Ser Ser Ser Ala Thr Phe Thr
              75              80              85

gct caa gaa caa ctt caa cca cca aat ctt gat ccg aaa gat gag atc 342
Ala Gln Glu Gln Leu Gln Pro Pro Asn Leu Asp Pro Lys Asp Glu Ile
      90              95              100

aac gtg ctt aag caa gag att gag atg ctt cag aaa ggg ata agc tat 390
Asn Val Leu Lys Gln Glu Ile Glu Met Leu Gln Lys Gly Ile Ser Tyr
      105              110              115

atg ttt gga gga gga gat ggg gct atg aat ctt gaa gaa ctt ctt ttg 438
Met Phe Gly Gly Gly Asp Gly Ala Met Asn Leu Glu Glu Leu Leu Leu
      120              125              130

ctt gag aag cat ctt gag tat tgg att tct cag att cgc tct gct aag 486
Leu Glu Lys His Leu Glu Tyr Trp Ile Ser Gln Ile Arg Ser Ala Lys
      135              140              145              150

atg gat gtt atg ctt caa gaa att cag tca ttg agg aac aag gaa gga 534
Met Asp Val Met Leu Gln Glu Ile Gln Ser Leu Arg Asn Lys Glu Gly
              155              160              165

gtc ctc aaa aac acc aac aag tat ctc ctc gac aag ata gag gaa aac 582
Val Leu Lys Asn Thr Asn Lys Tyr Leu Leu Asp Lys Ile Glu Glu Asn
      170              175              180

aac aat agc ata tta gat gct aac ttc gca gtc atg gag aca aac tat 630
Asn Asn Ser Ile Leu Asp Ala Asn Phe Ala Val Met Glu Thr Asn Tyr
      185              190              195

tcc tat ccg cta aca atg cca agt gaa ata ttt cag ttc tag 672
Ser Tyr Pro Leu Thr Met Pro Ser Glu Ile Phe Gln Phe *
```



200

205

210

accatagggt atttgaagac tatgtctcac gaatttaaata aaccttggtta agtataatat 732  
 agtggttgta aatcacacat aattaaaata aagcctgtgg aacttcgcta ggcagttgaa 792  
 aatctatccg tatgttttat cctcttggtt tacatttggtt ggtgtgaaga tgaaatgact 852  
 gcaagtgtgg tgtgtactta taactctttc tactttctat ctatgttttg aatttatgga 912  
 tt 914

&lt;210&gt; 232

&lt;211&gt; 211

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (2)...(57)

&lt;223&gt; Conserved domain

&lt;400&gt; 232

Met	Ala	Arg	Gly	Lys	Ile	Gln	Leu	Lys	Arg	Ile	Glu	Asn	Pro	Val	His
1				5				10					15		
Arg	Gln	Val	Thr	Phe	Cys	Lys	Arg	Arg	Thr	Gly	Leu	Leu	Lys	Lys	Ala
		20					25					30			
Lys	Glu	Leu	Ser	Val	Leu	Cys	Asp	Ala	Glu	Ile	Gly	Val	Val	Ile	Phe
	35					40					45				
Ser	Pro	Gln	Gly	Lys	Leu	Phe	Glu	Leu	Ala	Thr	Lys	Gly	Thr	Met	Glu
	50				55					60					
Gly	Met	Ile	Asp	Lys	Tyr	Met	Lys	Cys	Thr	Gly	Gly	Gly	Arg	Gly	Ser
65				70				75					80		
Ser	Ser	Ala	Thr	Phe	Thr	Ala	Gln	Glu	Gln	Leu	Gln	Pro	Pro	Asn	Leu
		85				90							95		
Asp	Pro	Lys	Asp	Glu	Ile	Asn	Val	Leu	Lys	Gln	Glu	Ile	Glu	Met	Leu
	100					105						110			
Gln	Lys	Gly	Ile	Ser	Tyr	Met	Phe	Gly	Gly	Gly	Asp	Gly	Ala	Met	Asn
	115				120						125				
Leu	Glu	Glu	Leu	Leu	Leu	Leu	Glu	Lys	His	Leu	Glu	Tyr	Trp	Ile	Ser
	130				135					140					
Gln	Ile	Arg	Ser	Ala	Lys	Met	Asp	Val	Met	Leu	Gln	Glu	Ile	Gln	Ser
145				150				155					160		
Leu	Arg	Asn	Lys	Glu	Gly	Val	Leu	Lys	Asn	Thr	Asn	Lys	Tyr	Leu	Leu
		165						170					175		
Asp	Lys	Ile	Glu	Glu	Asn	Asn	Asn	Ser	Ile	Leu	Asp	Ala	Asn	Phe	Ala
	180				185							190			
Val	Met	Glu	Thr	Asn	Tyr	Ser	Tyr	Pro	Leu	Thr	Met	Pro	Ser	Glu	Ile
	195				200						205				
Phe	Gln	Phe													
	210														

&lt;210&gt; 233

&lt;211&gt; 1276

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (47)...(1150)

&lt;400&gt; 233

```

taatccgatt cgtcttcatac tgattccctc ccttccgaga ataata atg tac ccg      55
                                     Met Tyr Pro
                                     1

cca cct ccc tca agc atc tac gct cct ccg atg ctg gtg aat tgc tcc      103
Pro Pro Pro Ser Ser Ile Tyr Ala Pro Pro Met Leu Val Asn Cys Ser
      5                      10                      15

ggt tgc cgg acg cct ctc cag ctc cca tcc ggc gcc cga tct att cgc      151
Gly Cys Arg Thr Pro Leu Gln Leu Pro Ser Gly Ala Arg Ser Ile Arg
      20                      25                      30                      35

tgc gct ctc tgc cag gct gtt act cat atc gcc gac cct cgc acc gcc      199
Cys Ala Leu Cys Gln Ala Val Thr His Ile Ala Asp Pro Arg Thr Ala
                      40                      45                      50

cct cct ccg caa cct tcc tcc gcc cct tct ccg cct ccc caa atc cac      247
Pro Pro Pro Gln Pro Ser Ser Ala Pro Ser Pro Pro Pro Gln Ile His
                      55                      60                      65

gcg cct ccc ggt cag ctg cct cac ccc cat ggc agg aag agg gcc gtg      295
Ala Pro Pro Gly Gln Leu Pro His Pro His Gly Arg Lys Arg Ala Val
                      70                      75                      80

atc tgt ggc atc tcg tat cgt ttc tct cgc cac gag ctc aaa ggc tgc      343
Ile Cys Gly Ile Ser Tyr Arg Phe Ser Arg His Glu Leu Lys Gly Cys
      85                      90                      95

atc aac gac gcc aag tgc atg cgt cac ctt ctc atc aac aaa ttc aaa      391
Ile Asn Asp Ala Lys Cys Met Arg His Leu Leu Ile Asn Lys Phe Lys
      100                      105                      110                      115

ttc tcc cca gat tca att ctc atg ctt acc gag gaa gaa act gat cca      439
Phe Ser Pro Asp Ser Ile Leu Met Leu Thr Glu Glu Glu Thr Asp Pro
                      120                      125                      130

tat cgt atc ccg acc aag caa aac atg agg atg gca ttg tat tgg ctc      487
Tyr Arg Ile Pro Thr Lys Gln Asn Met Arg Met Ala Leu Tyr Trp Leu
                      135                      140                      145

gta cag gga tgc aca gca ggc gac tca ctt gtc ttc cac tac tct ggt      535
Val Gln Gly Cys Thr Ala Gly Asp Ser Leu Val Phe His Tyr Ser Gly
                      150                      155                      160

cat ggt tcg cgt caa aga aac tac aac ggt gat gaa gtt gat ggc tat      583
His Gly Ser Arg Gln Arg Asn Tyr Asn Gly Asp Glu Val Asp Gly Tyr
                      165                      170                      175

gat gaa aca ctc tgt cct ctg gat ttt gaa act cag ggg atg att gta      631
Asp Glu Thr Leu Cys Pro Leu Asp Phe Glu Thr Gln Gly Met Ile Val
      180                      185                      190                      195

gac gat gag atc aac gca acc att gta cgc cct ctt cca cat ggt gtc      679
Asp Asp Glu Ile Asn Ala Thr Ile Val Arg Pro Leu Pro His Gly Val
                      200                      205                      210

```

```

aag ctc cat tca att atc gat gct tgc cat agt ggt acc gtt ctg gat 727
Lys Leu His Ser Ile Ile Asp Ala Cys His Ser Gly Thr Val Leu Asp
      215              220              225

tta ccc ttc cta tgc aga atg aac aga gct ggg cag tat gtg tgg gag 775
Leu Pro Phe Leu Cys Arg Met Asn Arg Ala Gly Gln Tyr Val Trp Glu
      230              235              240

gat cat cgg cct agg tca ggt ttg tgg aaa gga act gct ggt gga gaa 823
Asp His Arg Pro Arg Ser Gly Leu Trp Lys Gly Thr Ala Gly Gly Glu
      245              250              255

gcc att tca att agt gga tgt gat gat gat cag act tcg gcc gac aca 871
Ala Ile Ser Ile Ser Gly Cys Asp Asp Asp Gln Thr Ser Ala Asp Thr
      260              265              270              275

tca gcg ctg tcg aag atc acg tct acg ggt gct atg act ttc tgt ttt 919
Ser Ala Leu Ser Lys Ile Thr Ser Thr Gly Ala Met Thr Phe Cys Phe
      280              285              290

att caa gca att gaa cgc agc gca caa ggc aca acc tat gga agc ctt 967
Ile Gln Ala Ile Glu Arg Ser Ala Gln Gly Thr Thr Tyr Gly Ser Leu
      295              300              305

ctg aat tct atg cgc acc aca ata agg aat aca ggg aat gat ggt ggt 1015
Leu Asn Ser Met Arg Thr Thr Ile Arg Asn Thr Gly Asn Asp Gly Gly
      310              315              320

ggt agt ggt gga gtt gtg acg act gtg ctg agc atg ctt ctg aca ggg 1063
Gly Ser Gly Gly Val Val Thr Thr Val Leu Ser Met Leu Leu Thr Gly
      325              330              335

gga agt gcg att ggg gga tta aga cag gag cct caa ctg act gct tgc 1111
Gly Ser Ala Ile Gly Gly Leu Arg Gln Glu Pro Gln Leu Thr Ala Cys
      340              345              350              355

caa aca ttc gat gtc tat gca aag cct ttc act ctc tag taaaggacaa 1160
Gln Thr Phe Asp Val Tyr Ala Lys Pro Phe Thr Leu *
      360              365

gtcacttttt atgtatagcg agtgtgattt gagaatccgt ccatataacc accttttgtt 1220
tcttattttt atttttcttt caaaagaata aaggaaaaca ttgatttggt gattcg 1276

<210> 234
<211> 367
<212> PRT
<213> Arabidopsis thaliana

<220>
<221> DOMAIN
<222> (18)...(39)
<223> Conserved domain

<400> 234
Met Tyr Pro Pro Pro Pro Ser Ser Ile Tyr Ala Pro Pro Met Leu Val
  1              5              10              15

```

```

Asn Cys Ser Gly Cys Arg Thr Pro Leu Gln Leu Pro Ser Gly Ala Ar
      20      25      30
Ser Ile Arg Cys Ala Leu Cys Gln Ala Val Thr His Ile Ala Asp Pro
      35      40      45
Arg Thr Ala Pro Pro Pro Gln Pro Ser Ser Ala Pro Ser Pro Pro Pro
      50      55      60
Gln Ile His Ala Pro Pro Gly Gln Leu Pro His Pro His Gly Arg Lys
      65      70      75      80
Arg Ala Val Ile Cys Gly Ile Ser Tyr Arg Phe Ser Arg His Glu Leu
      85      90      95
Lys Gly Cys Ile Asn Asp Ala Lys Cys Met Arg His Leu Leu Ile Asn
      100      105      110
Lys Phe Lys Phe Ser Pro Asp Ser Ile Leu Met Leu Thr Glu Glu Glu
      115      120      125
Thr Asp Pro Tyr Arg Ile Pro Thr Lys Gln Asn Met Arg Met Ala Leu
      130      135      140
Tyr Trp Leu Val Gln Gly Cys Thr Ala Gly Asp Ser Leu Val Phe His
      145      150      155      160
Tyr Ser Gly His Gly Ser Arg Gln Arg Asn Tyr Asn Gly Asp Glu Val
      165      170      175
Asp Gly Tyr Asp Glu Thr Leu Cys Pro Leu Asp Phe Glu Thr Gln Gly
      180      185      190
Met Ile Val Asp Asp Glu Ile Asn Ala Thr Ile Val Arg Pro Leu Pro
      195      200      205
His Gly Val Lys Leu His Ser Ile Ile Asp Ala Cys His Ser Gly Thr
      210      215      220
Val Leu Asp Leu Pro Phe Leu Cys Arg Met Asn Arg Ala Gly Gln Tyr
      225      230      235      240
Val Trp Glu Asp His Arg Pro Arg Ser Gly Leu Trp Lys Gly Thr Ala
      245      250      255
Gly Gly Glu Ala Ile Ser Ile Ser Gly Cys Asp Asp Asp Gln Thr Ser
      260      265      270
Ala Asp Thr Ser Ala Leu Ser Lys Ile Thr Ser Thr Gly Ala Met Thr
      275      280      285
Phe Cys Phe Ile Gln Ala Ile Glu Arg Ser Ala Gln Gly Thr Thr Tyr
      290      295      300
Gly Ser Leu Leu Asn Ser Met Arg Thr Thr Ile Arg Asn Thr Gly Asn
      305      310      315      320
Asp Gly Gly Gly Ser Gly Gly Val Val Thr Thr Val Leu Ser Met Leu
      325      330      335
Leu Thr Gly Gly Ser Ala Ile Gly Gly Leu Arg Gln Glu Pro Gln Leu
      340      345      350
Thr Ala Cys Gln Thr Phe Asp Val Tyr Ala Lys Pro Phe Thr Leu
      355      360      365

```

&lt;210&gt; 235

&lt;211&gt; 2545

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (103)...(2322)

&lt;400&gt; 235

acatatgttt taaattcttt gtctgaatct tacaggatcc gagagagaga gctctggaac 60

gatattaaca tatatcatga agaaaaagat tgaagtattg at atg gga ata act  
Met Gly Ile Thr  
1

aaa act tct cct aat act aca att ctc ttg aag act ttt cac aat aat 162  
Lys Thr Ser Pro Asn Thr Thr Ile Leu Leu Lys Thr Phe His Asn Asn  
5 10 15 20

tct atg tcc caa gat tat cat cat cat cat cat cat aat caa cac caa 210  
Ser Met Ser Gln Asp Tyr His His His His His His Asn Gln His Gln  
25 30 35

gga ggt atc ttc aac ttc tct aat gga ttc gac cga tca gat tct ccc 258  
Gly Gly Ile Phe Asn Phe Ser Asn Gly Phe Asp Arg Ser Asp Ser Pro  
40 45 50

aat tta aca act cag cag aag caa gag cat caa agg gta gag atg gac 306  
Asn Leu Thr Thr Gln Gln Lys Gln Glu His Gln Arg Val Glu Met Asp  
55 60 65

gag gaa tct tca gtc gcc gga ggt agg att ccg gtc tac gaa tca gcc 354  
Glu Glu Ser Ser Val Ala Gly Gly Arg Ile Pro Val Tyr Glu Ser Ala  
70 75 80

ggt atg tta tcc gaa atg ttt aat ttc ccc gga agc agc ggt gga gga 402  
Gly Met Leu Ser Glu Met Phe Asn Phe Pro Gly Ser Ser Gly Gly Gly  
85 90 95 100

aga gat ctc gac ctc ggc caa tct ttc cgg tca aat agg cag ttg ctt 450  
Arg Asp Leu Asp Leu Gly Gln Ser Phe Arg Ser Asn Arg Gln Leu Leu  
105 110 115

gag gag caa cat cag aat att ccg gct atg aat gct acg gat tca gcc 498  
Glu Glu Gln His Gln Asn Ile Pro Ala Met Asn Ala Thr Asp Ser Ala  
120 125 130

acc gcc acc gca gcc gcc atg cag tta ttc ttg atg aat cca ccg cca 546  
Thr Ala Thr Ala Ala Ala Met Gln Leu Phe Leu Met Asn Pro Pro Pro  
135 140 145

ccg caa caa cca ccg tct ccg tca tcc aca act tcc cca agg agc cac 594  
Pro Gln Gln Pro Pro Ser Pro Ser Ser Thr Thr Ser Pro Arg Ser His  
150 155 160

cac aat tct tca act ctt cac atg tta ctt cca agt cca tcc acc aac 642  
His Asn Ser Ser Thr Leu His Met Leu Leu Pro Ser Pro Ser Thr Asn  
165 170 175 180

aca act cac cat cag aac tac act aat cat atg tct atg cat cag ctt 690  
Thr Thr His His Gln Asn Tyr Thr Asn His Met Ser Met His Gln Leu  
185 190 195

cca cat cag cat cac caa cag ata tcg acg tgg cag tct tct ccc gat 738  
Pro His Gln His His Gln Gln Ile Ser Thr Trp Gln Ser Ser Pro Asp  
200 205 210

cat cat cat cat cat cac aac agc caa acg gag att ggg acc gtc cac 786

His	His	His	His	His	His	Asn	Ser	Gln	Thr	Glu	Ile	Gly	Thr	Val	Hi	
		215					220					225				
gtg	gaa	aac	agc	gga	gga	cac	gga	gga	caa	ggc	ttg	tcc	tta	tct	ctc	834
Val	Glu	Asn	Ser	Gly	Gly	His	Gly	Gly	Gln	Gly	Leu	Ser	Leu	Ser	Leu	
		230				235					240					
tca	tcg	tct	tta	gag	gct	gca	gca	aaa	gcg	gaa	gag	tat	aga	aac	att	882
Ser	Ser	Ser	Leu	Glu	Ala	Ala	Ala	Lys	Ala	Glu	Glu	Tyr	Arg	Asn	Ile	
		245			250					255					260	
tac	tac	gga	gcc	aat	tct	tct	aac	gca	tca	cct	cat	cat	caa	tac	aat	930
Tyr	Tyr	Gly	Ala	Asn	Ser	Ser	Asn	Ala	Ser	Pro	His	His	Gln	Tyr	Asn	
				265					270						275	
caa	ttc	aag	act	ctt	ctt	gct	aat	tct	tct	caa	cat	cac	cat	caa	gta	978
Gln	Phe	Lys	Thr	Leu	Leu	Ala	Asn	Ser	Ser	Gln	His	His	His	Gln	Val	
			280					285						290		
tta	aac	caa	ttc	cga	tca	tct	ccg	gct	gct	tct	tcc	tct	tcc	atg	gca	1026
Leu	Asn	Gln	Phe	Arg	Ser	Ser	Pro	Ala	Ala	Ser	Ser	Ser	Ser	Met	Ala	
		295					300					305				
gcg	gtc	aat	atc	tta	aga	aac	tcg	agg	tac	aca	acg	gcc	gcg	caa	gag	1074
Ala	Val	Asn	Ile	Leu	Arg	Asn	Ser	Arg	Tyr	Thr	Thr	Ala	Ala	Gln	Glu	
		310				315						320				
ttg	ttg	gaa	gag	ttt	tgt	agt	gtt	gga	aga	gga	ttt	ttg	aag	aag	aac	1122
Leu	Leu	Glu	Glu	Phe	Cys	Ser	Val	Gly	Arg	Gly	Phe	Leu	Lys	Lys	Asn	
		325			330					335					340	
aaa	ctt	ggg	aac	agc	tca	aac	cct	aat	act	tgc	ggg	ggg	gat	ggg	ggg	1170
Lys	Leu	Gly	Asn	Ser	Ser	Asn	Pro	Asn	Thr	Cys	Gly	Gly	Asp	Gly	Gly	
			345						350					355		
ggc	agc	tct	cct	tcg	tcg	gcc	gga	gca	aac	aag	gag	cat	cct	cct	tta	1218
Gly	Ser	Ser	Pro	Ser	Ser	Ala	Gly	Ala	Asn	Lys	Glu	His	Pro	Pro	Leu	
			360					365					370			
tcg	gcg	tct	gat	cgg	att	gag	cat	caa	aga	agg	aaa	gtg	aaa	cta	ctc	1266
Ser	Ala	Ser	Asp	Arg	Ile	Glu	His	Gln	Arg	Arg	Lys	Val	Lys	Leu	Leu	
		375					380					385				
acc	atg	ctt	gaa	gag	gtg	gac	cga	cgg	tac	aac	cat	tac	tgc	gag	caa	1314
Thr	Met	Leu	Glu	Glu	Val	Asp	Arg	Arg	Tyr	Asn	His	Tyr	Cys	Glu	Gln	
		390				395					400					
atg	cag	atg	gtt	gtg	aac	tct	ttc	gac	ata	gta	atg	ggc	cac	ggg	gcg	1362
Met	Gln	Met	Val	Val	Asn	Ser	Phe	Asp	Ile	Val	Met	Gly	His	Gly	Ala	
		405			410					415					420	
gca	tta	ccg	tac	acc	gca	ttg	gct	caa	aaa	gct	atg	tca	aga	cat	ttt	1410
Ala	Leu	Pro	Tyr	Thr	Ala	Leu	Ala	Gln	Lys	Ala	Met	Ser	Arg	His	Phe	
			425					430					435			
aga	tgc	ctt	aaa	gat	gca	gtt	gcg	gct	cag	ctt	aag	cag	agt	tgc	gaa	1458
Arg	Cys	Leu	Lys	Asp	Ala	Val	Ala	Ala	Gln	Leu	Lys	Gln	Ser	Cys	Glu	

440	445	450	
ctt ctt ggg gac aaa gat gca gcg gga atc tct tct tcc ggg tta aca			1506
Leu Leu Gly Asp Lys Asp Ala Ala Gly Ile Ser Ser Ser Gly Leu Thr			
455	460	465	
aaa ggt gaa act ccg cgt ttg cgt ttg cta gag caa agt ttg cgt cag			1554
Lys Gly Glu Thr Pro Arg Leu Arg Leu Leu Glu Gln Ser Leu Arg Gln			
470	475	480	
caa cgt gcg ttt cat caa atg ggt atg atg gaa caa gaa gct tgg cgg			1602
Gln Arg Ala Phe His Gln Met Gly Met Met Glu Gln Glu Ala Trp Arg			
485	490	495	500
cca caa cgc ggt ttg cct gaa cgc tcc gtc aat ata ctt aga gct tgg			1650
Pro Gln Arg Gly Leu Pro Glu Arg Ser Val Asn Ile Leu Arg Ala Trp			
505	510	515	
ctc ttc gaa cat ttc ctt cac ccg tat cca agt gat gca gat aaa cac			1698
Leu Phe Glu His Phe Leu His Pro Tyr Pro Ser Asp Ala Asp Lys His			
520	525	530	
cta ttg gct cga cag act ggt tta tcc aga aat cag gta tca aat tgg			1746
Leu Leu Ala Arg Gln Thr Gly Leu Ser Arg Asn Gln Val Ser Asn Trp			
535	540	545	
ttc ata aat gct agg gtt cgt tta tgg aaa cca atg gtg gaa gaa atg			1794
Phe Ile Asn Ala Arg Val Arg Leu Trp Lys Pro Met Val Glu Glu Met			
550	555	560	
tac caa caa gaa tca aaa gaa aga gaa aga gaa gag gaa tta gaa gag			1842
Tyr Gln Gln Glu Ser Lys Glu Arg Glu Arg Glu Glu Glu Leu Glu Glu			
565	570	575	580
aac gaa gaa gat caa gaa aca aaa aac agc aac gac gac aag agc aca			1890
Asn Glu Glu Asp Gln Glu Thr Lys Asn Ser Asn Asp Asp Lys Ser Thr			
585	590	595	
aaa tcc aac aac aat gaa agc aac ttc act gcc gtt cgg acc act tca			1938
Lys Ser Asn Asn Asn Glu Ser Asn Phe Thr Ala Val Arg Thr Thr Ser			
600	605	610	
caa act cca acg aca acc gca cca gac gca tca gac gca gac gca gca			1986
Gln Thr Pro Thr Thr Thr Ala Pro Asp Ala Ser Asp Ala Asp Ala Ala			
615	620	625	
gta gcg aca ggc cac cgt cta aga tcc aac att aat gct tac gaa aac			2034
Val Ala Thr Gly His Arg Leu Arg Ser Asn Ile Asn Ala Tyr Glu Asn			
630	635	640	
gac gct tca tca ctt cta ctc cct tcc tct tat tcc aac gcc gcc gct			2082
Asp Ala Ser Ser Leu Leu Leu Pro Ser Ser Tyr Ser Asn Ala Ala Ala			
645	650	655	660
cct gcc gct gtt tct gac gac ttg aat tct cgt tac ggt ggc tca gac			2130
Pro Ala Ala Val Ser Asp Asp Leu Asn Ser Arg Tyr Gly Gly Ser Asp			
665	670	675	

```

gcg ttt tcc gcc gtt gcc acg tgt caa caa agt gta ggt ggg ttc gat 2178
Ala Phe Ser Ala Val Ala Thr Cys Gln Gln Ser Val Gly Gly Phe Asp
      680                      685                      690

gat gct gac atg gat ggt gtt aac gtt ata agg ttt ggg aca aac cct 2226
Asp Ala Asp Met Asp Gly Val Asn Val Ile Arg Phe Gly Thr Asn Pro
      695                      700                      705

act ggt gac gtg tct ctc acg ctt ggt tta cgc cac gct gga aac atg 2274
Thr Gly Asp Val Ser Leu Thr Leu Gly Leu Arg His Ala Gly Asn Met
      710                      715                      720

cct gac aaa gac gct tct ttc tgc gtt aga gag ttt ggg ggt ttt tag 2322
Pro Asp Lys Asp Ala Ser Phe Cys Val Arg Glu Phe Gly Gly Phe *
      725                      730                      735

tttgc ttttg tca ctccatt taattaatta attatagttt tccattctta cttatttttaa 2382
ttgaaaatct atttttgtct cttaaaagtc caaacaatac attagtctag ccctcctctg 2442
cttttttttt tctatctcgt gaagagaaga aaacgatacg taaatccctt cgaaaactaa 2502
tgtacgttgt acgacttatt gttttcataa aaaaaaaaaa aaa 2545

```

&lt;210&gt; 236

&lt;211&gt; 739

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN .

&lt;222&gt; (500)...(560)

&lt;223&gt; Conserved domain

&lt;400&gt; 236

```

Met Gly Ile Thr Lys Thr Ser Pro Asn Thr Thr Ile Leu Leu Lys Thr
 1          5          10          15
Phe His Asn Asn Ser Met Ser Gln Asp Tyr His His His His His
      20          25          30
Asn Gln His Gln Gly Gly Ile Phe Asn Phe Ser Asn Gly Phe Asp Arg
      35          40          45
Ser Asp Ser Pro Asn Leu Thr Thr Gln Gln Lys Gln Glu His Gln Arg
      50          55          60
Val Glu Met Asp Glu Glu Ser Ser Val Ala Gly Gly Arg Ile Pro Val
      65          70          75          80
Tyr Glu Ser Ala Gly Met Leu Ser Glu Met Phe Asn Phe Pro Gly Ser
      85          90          95
Ser Gly Gly Gly Arg Asp Leu Asp Leu Gly Gln Ser Phe Arg Ser Asn
      100          105          110
Arg Gln Leu Leu Glu Glu Gln His Gln Asn Ile Pro Ala Met Asn Ala
      115          120          125
Thr Asp Ser Ala Thr Ala Thr Ala Ala Met Gln Leu Phe Leu Met
      130          135          140
Asn Pro Pro Pro Pro Gln Gln Pro Pro Ser Pro Ser Ser Thr Thr Ser
      145          150          155          160
Pro Arg Ser His His Asn Ser Ser Thr Leu His Met Leu Leu Pro Ser
      165          170          175
Pro Ser Thr Asn Thr Thr His His Gln Asn Tyr Thr Asn His Met Ser
      180          185          190

```



Met His Gln Leu Pro His Gln His His Gln Gln Ile Ser Thr Trp Gl  
 195 200 205  
 Ser Ser Pro Asp His His His His His His Asn Ser Gln Thr Glu Ile  
 210 215 220  
 Gly Thr Val His Val Glu Asn Ser Gly Gly His Gly Gly Gln Gly Leu  
 225 230 235 240  
 Ser Leu Ser Leu Ser Ser Ser Leu Glu Ala Ala Ala Lys Ala Glu Glu  
 245 250 255  
 Tyr Arg Asn Ile Tyr Tyr Gly Ala Asn Ser Ser Asn Ala Ser Pro His  
 260 265 270  
 His Gln Tyr Asn Gln Phe Lys Thr Leu Leu Ala Asn Ser Ser Gln His  
 275 280 285  
 His His Gln Val Leu Asn Gln Phe Arg Ser Ser Pro Ala Ala Ser Ser  
 290 295 300  
 Ser Ser Met Ala Ala Val Asn Ile Leu Arg Asn Ser Arg Tyr Thr Thr  
 305 310 315 320  
 Ala Ala Gln Glu Leu Leu Glu Glu Phe Cys Ser Val Gly Arg Gly Phe  
 325 330 335  
 Leu Lys Lys Asn Lys Leu Gly Asn Ser Ser Asn Pro Asn Thr Cys Gly  
 340 345 350  
 Gly Asp Gly Gly Gly Ser Ser Pro Ser Ser Ala Gly Ala Asn Lys Glu  
 355 360 365  
 His Pro Pro Leu Ser Ala Ser Asp Arg Ile Glu His Gln Arg Arg Lys  
 370 375 380  
 Val Lys Leu Leu Thr Met Leu Glu Glu Val Asp Arg Arg Tyr Asn His  
 385 390 395 400  
 Tyr Cys Glu Gln Met Gln Met Val Val Asn Ser Phe Asp Ile Val Met  
 405 410 415  
 Gly His Gly Ala Ala Leu Pro Tyr Thr Ala Leu Ala Gln Lys Ala Met  
 420 425 430  
 Ser Arg His Phe Arg Cys Leu Lys Asp Ala Val Ala Ala Gln Leu Lys  
 435 440 445  
 Gln Ser Cys Glu Leu Leu Gly Asp Lys Asp Ala Ala Gly Ile Ser Ser  
 450 455 460  
 Ser Gly Leu Thr Lys Gly Glu Thr Pro Arg Leu Arg Leu Leu Glu Gln  
 465 470 475 480  
 Ser Leu Arg Gln Gln Arg Ala Phe His Gln Met Gly Met Met Glu Gln  
 485 490 495  
 Glu Ala Trp Arg Pro Gln Arg Gly Leu Pro Glu Arg Ser Val Asn Ile  
 500 505 510  
 Leu Arg Ala Trp Leu Phe Glu His Phe Leu His Pro Tyr Pro Ser Asp  
 515 520 525  
 Ala Asp Lys His Leu Leu Ala Arg Gln Thr Gly Leu Ser Arg Asn Gln  
 530 535 540  
 Val Ser Asn Trp Phe Ile Asn Ala Arg Val Arg Leu Trp Lys Pro Met  
 545 550 555 560  
 Val Glu Glu Met Tyr Gln Gln Glu Ser Lys Glu Arg Glu Arg Glu Glu  
 565 570 575  
 Glu Leu Glu Glu Asn Glu Glu Asp Gln Glu Thr Lys Asn Ser Asn Asp  
 580 585 590  
 Asp Lys Ser Thr Lys Ser Asn Asn Asn Glu Ser Asn Phe Thr Ala Val  
 595 600 605  
 Arg Thr Thr Ser Gln Thr Pro Thr Thr Thr Ala Pro Asp Ala Ser Asp  
 610 615 620  
 Ala Asp Ala Ala Val Ala Thr Gly His Arg Leu Arg Ser Asn Ile Asn  
 625 630 635 640  
 Ala Tyr Glu Asn Asp Ala Ser Ser Leu Leu Leu Pro Ser Ser Tyr Ser

```

        645              650              655
Asn Ala Ala Ala Pro Ala Ala Val Ser Asp Asp Leu Asn Ser Arg Tyr
        660              665              670
Gly Gly Ser Asp Ala Phe Ser Ala Val Ala Thr Cys Gln Gln Ser Val
        675              680              685
Gly Gly Phe Asp Asp Ala Asp Met Asp Gly Val Asn Val Ile Arg Phe
        690              695              700
Gly Thr Asn Pro Thr Gly Asp Val Ser Leu Thr Leu Gly Leu Arg His
705              710              715              720
Ala Gly Asn Met Pro Asp Lys Asp Ala Ser Phe Cys Val Arg Glu Phe
        725              730              735
Gly Gly Phe

```

&lt;210&gt; 237

&lt;211&gt; 1179

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (109)...(966)

&lt;400&gt; 237

```

cttctctctt ctcaaaaacc cttccctctt cgtctccaaa caacaacaaa cacaacaaca 60
acaaaaatct tacaagaaga tcatttttag aaaccctatt aggataaa atg gat tac 117
                                     Met Asp Tyr
                                     1

```

```

gag gca tca aga atc gtc gaa atg gta gaa gat gaa gaa cat ata gat 165
Glu Ala Ser Arg Ile Val Glu Met Val Glu Asp Glu Glu His Ile Asp
    5              10              15

```

```

cta cca cca gga ttc aga ttt cac cct act gat gaa gaa ctc ata act 213
Leu Pro Pro Gly Phe Arg Phe His Pro Thr Asp Glu Glu Leu Ile Thr
    20              25              30              35

```

```

cac tac ctc aaa cca aag gtt ttc aac act ttc ttc tct gct act gcc 261
His Tyr Leu Lys Pro Lys Val Phe Asn Thr Phe Phe Ser Ala Thr Ala
        40              45              50

```

```

att ggt gaa gtt gat ctc aac aag att gag cct tgg gac tta cca tgg 309
Ile Gly Glu Val Asp Leu Asn Lys Ile Glu Pro Trp Asp Leu Pro Trp
    55              60              65

```

```

aag gct aag atg gga gaa aaa gaa tgg tat ttc ttc tgt gtg aga gac 357
Lys Ala Lys Met Gly Glu Lys Glu Trp Tyr Phe Phe Cys Val Arg Asp
    70              75              80

```

```

cgg aaa tac ccg acc ggt tta agg aca aac cgg gcg aca gaa gcc ggt 405
Arg Lys Tyr Pro Thr Gly Leu Arg Thr Asn Arg Ala Thr Glu Ala Gly
    85              90              95

```

```

tat tgg aaa gcc aca gga aaa gac aaa gag ata ttc aag gga aaa tca 453
Tyr Trp Lys Ala Thr Gly Lys Asp Lys Glu Ile Phe Lys Gly Lys Ser
100              105              110              115

```

```

ctt gtg ggt atg aag aaa act ttg gtt ttc tat aaa gga aga gct cct 501
Leu Val Gly Met Lys Lys Thr Leu Val Phe Tyr Lys Gly Arg Ala Pro
120 125 130

aaa gga gtt aaa acc aat tgg gtt atg cat gaa tat cgt tta gaa ggc 549
Lys Gly Val Lys Thr Asn Trp Val Met His Glu Tyr Arg Leu Glu Gly
135 140 145

aaa tat tgt att gaa aat ctt ccc caa aca gct aag aac gaa tgg gtt 597
Lys Tyr Cys Ile Glu Asn Leu Pro Gln Thr Ala Lys Asn Glu Trp Val
150 155 160

ata tgt cgt gtt ttc caa aaa cgt gcc gat ggt aca aag gtt cca atg 645
Ile Cys Arg Val Phe Gln Lys Arg Ala Asp Gly Thr Lys Val Pro Met
165 170 175

tca atg ctt gat cca cac att aac cga atg gaa cca gcc ggt tta cct 693
Ser Met Leu Asp Pro His Ile Asn Arg Met Glu Pro Ala Gly Leu Pro
180 185 190 195

tcg tta atg gat tgt tct caa cga gac tcc ttc acc ggt tcg tcg tct 741
Ser Leu Met Asp Cys Ser Gln Arg Asp Ser Phe Thr Gly Ser Ser Ser
200 205 210

cac gtg acc tgc ttc tcc gac caa gaa acc gaa gac aaa aga ctt gtc 789
His Val Thr Cys Phe Ser Asp Gln Glu Thr Glu Asp Lys Arg Leu Val
215 220 225

cac gag tcc aaa gac ggt ttt ggt tct ctg ttt tac tcg gat cct ctg 837
His Glu Ser Lys Asp Gly Phe Gly Ser Leu Phe Tyr Ser Asp Pro Leu
230 235 240

ttt tta caa gac aat tat tcg cta atg aag ctg ttg ctt gac ggt caa 885
Phe Leu Gln Asp Asn Tyr Ser Leu Met Lys Leu Leu Leu Asp Gly Gln
245 250 255

gaa act caa ttc tcc ggc aaa cct ttc gac ggt cgt gat tcg tcc ggt 933
Glu Thr Gln Phe Ser Gly Lys Pro Phe Asp Gly Arg Asp Ser Ser Gly
260 265 270 275

aca gaa gaa ttg gat tgc gtt tgg aat ttc tga gttgtataag ttatgttgta 986
Thr Glu Glu Leu Asp Cys Val Trp Asn Phe *
280 285

gacttgtagt agtcatgtgt tcgtgtgtgt gaatgaatat tcttgttaca tttttttgta 1046
aaaaaggaga aaaaaatatg ctagaaagtc aattgctttt gttatgtagc attagtgttt 1106
tttatgtact caatagactt cctaattaaa taaaaatctt aatttatttg ccaaaaaaaaa 1166
aaaaaaaaaaa aaa 1179

<210> 238
<211> 285
<212> PRT
<213> Arabidopsis thaliana

<220>
<221> DOMAIN

```

&lt;222&gt; (23)...(167)

&lt;223&gt; Conserved domain

&lt;400&gt; 238

```

Met Asp Tyr Glu Ala Ser Arg Ile Val Glu Met Val Glu Asp Glu Glu
 1           5           10           15
His Ile Asp Leu Pro Pro Gly Phe Arg Phe His Pro Thr Asp Glu Glu
          20           25           30
Leu Ile Thr His Tyr Leu Lys Pro Lys Val Phe Asn Thr Phe Phe Ser
          35           40           45
Ala Thr Ala Ile Gly Glu Val Asp Leu Asn Lys Ile Glu Pro Trp Asp
 50           55           60
Leu Pro Trp Lys Ala Lys Met Gly Glu Lys Glu Trp Tyr Phe Phe Cys
65           70           75           80
Val Arg Asp Arg Lys Tyr Pro Thr Gly Leu Arg Thr Asn Arg Ala Thr
          85           90           95
Glu Ala Gly Tyr Trp Lys Ala Thr Gly Lys Asp Lys Glu Ile Phe Lys
          100          105          110
Gly Lys Ser Leu Val Gly Met Lys Lys Thr Leu Val Phe Tyr Lys Gly
          115          120          125
Arg Ala Pro Lys Gly Val Lys Thr Asn Trp Val Met His Glu Tyr Arg
          130          135          140
Leu Glu Gly Lys Tyr Cys Ile Glu Asn Leu Pro Gln Thr Ala Lys Asn
145          150          155          160
Glu Trp Val Ile Cys Arg Val Phe Gln Lys Arg Ala Asp Gly Thr Lys
          165          170          175
Val Pro Met Ser Met Leu Asp Pro His Ile Asn Arg Met Glu Pro Ala
          180          185          190
Gly Leu Pro Ser Leu Met Asp Cys Ser Gln Arg Asp Ser Phe Thr Gly
          195          200          205
Ser Ser Ser His Val Thr Cys Phe Ser Asp Gln Glu Thr Glu Asp Lys
          210          215          220
Arg Leu Val His Glu Ser Lys Asp Gly Phe Gly Ser Leu Phe Tyr Ser
225          230          235          240
Asp Pro Leu Phe Leu Gln Asp Asn Tyr Ser Leu Met Lys Leu Leu Leu
          245          250          255
Asp Gly Gln Glu Thr Gln Phe Ser Gly Lys Pro Phe Asp Gly Arg Asp
          260          265          270
Ser Ser Gly Thr Glu Glu Leu Asp Cys Val Trp Asn Phe
          275          280          285

```

&lt;210&gt; 239

&lt;211&gt; 1695

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(1695)

&lt;400&gt; 239

```

atg gtt gat tct tct cgt gat tcg tgt ttc aaa gct ggt aag ttt agt   48
Met Val Asp Ser Ser Arg Asp Ser Cys Phe Lys Ala Gly Lys Phe Ser
 1           5           10           15

gct cca ggt ttt cga ttt cac cct act gat gaa gag ctt gtg gtt tat   96

```

Ala	Pro	Gly	Phe	Arg	Phe	His	Pro	Thr	Asp	Glu	Glu	Leu	Val	Val	Ty	
			20					25					30			
tat	ctt	aag	agg	aag	atc	tgt	tgt	aaa	aaa	ctt	cga	gtc	aat	gcc	att	144
Tyr	Leu	Lys	Arg	Lys	Ile	Cys	Cys	Lys	Lys	Leu	Arg	Val	Asn	Ala	Ile	
		35				40					45					
ggt	gtc	gtt	gat	gtt	tac	aaa	gtc	gat	cct	tct	gaa	ttg	cct	ggt	cta	192
Gly	Val	Val	Asp	Val	Tyr	Lys	Val	Asp	Pro	Ser	Glu	Leu	Pro	Gly	Leu	
	50					55					60					
tcg	atg	ttg	aag	acg	gga	gat	aga	cag	tgg	ttc	ttt	ttc	act	cca	agg	240
Ser	Met	Leu	Lys	Thr	Gly	Asp	Arg	Gln	Trp	Phe	Phe	Phe	Thr	Pro	Arg	
	65				70				75						80	
aat	agg	aag	tat	cct	aac	gca	gct	agg	tca	agt	aga	ggt	act	gca	act	288
Asn	Arg	Lys	Tyr	Pro	Asn	Ala	Ala	Arg	Ser	Ser	Arg	Gly	Thr	Ala	Thr	
				85				90						95		
ggt	tat	tgg	aag	gcg	aca	gga	aag	gat	cga	gtc	att	gag	tac	aat	tca	336
Gly	Tyr	Trp	Lys	Ala	Thr	Gly	Lys	Asp	Arg	Val	Ile	Glu	Tyr	Asn	Ser	
			100					105					110			
aga	tct	gtt	gga	ctc	aag	aag	act	ctt	gtt	ttc	tat	aga	ggt	cgt	gct	384
Arg	Ser	Val	Gly	Leu	Lys	Lys	Thr	Leu	Val	Phe	Tyr	Arg	Gly	Arg	Ala	
		115					120					125				
cct	aat	ggt	gag	aga	act	gac	tgg	gtg	atg	cat	gag	tac	act	atg	gat	432
Pro	Asn	Gly	Glu	Arg	Thr	Asp	Trp	Val	Met	His	Glu	Tyr	Thr	Met	Asp	
	130					135					140					
gaa	gaa	gag	cta	ggg	aga	tgt	aag	aac	gct	aag	gag	tat	tat	gct	ctt	480
Glu	Glu	Glu	Leu	Gly	Arg	Cys	Lys	Asn	Ala	Lys	Glu	Tyr	Tyr	Ala	Leu	
	145				150					155				160		
tat	aag	ctt	tat	aag	aag	agt	ggg	gct	ggt	cct	aag	aat	ggt	gaa	cag	528
Tyr	Lys	Leu	Tyr	Lys	Lys	Ser	Gly	Ala	Gly	Pro	Lys	Asn	Gly	Glu	Gln	
				165				170						175		
tat	ggt	gct	ccg	ttc	caa	gaa	gaa	gaa	tgg	gtt	gat	agt	gat	agt	gaa	576
Tyr	Gly	Ala	Pro	Phe	Gln	Glu	Glu	Glu	Trp	Val	Asp	Ser	Asp	Ser	Glu	
			180					185					190			
gat	gca	gat	agt	gtc	gct	gta	ccg	gat	tat	ccc	gtg	gtc	cgt	tat	gag	624
Asp	Ala	Asp	Ser	Val	Ala	Val	Pro	Asp	Tyr	Pro	Val	Val	Arg	Tyr	Glu	
		195					200					205				
aat	ggt	cct	tgt	gtg	gat	gat	act	aaa	ttt	tgc	aat	cct	gtc	aaa	ctt	672
Asn	Gly	Pro	Cys	Val	Asp	Asp	Thr	Lys	Phe	Cys	Asn	Pro	Val	Lys	Leu	
	210					215					220					
cag	tta	gag	gat	att	gag	aag	ctt	ctc	aat	gaa	atc	cca	gat	gca	ccc	720
Gln	Leu	Glu	Asp	Ile	Glu	Lys	Leu	Leu	Asn	Glu	Ile	Pro	Asp	Ala	Pro	
	225				230				235					240		
ggg	gtt	aac	caa	aga	cag	ttt	gat	gag	ttt	gtt	ggt	gtt	cca	cag	ggt	768
Gly	Val	Asn	Gln	Arg	Gln	Phe	Asp	Glu	Phe	Val	Gly	Val	Pro	Gln	Gly	

245	250	255	
aat agt gca gaa gtg ata cag agc aca ttg ctg aat aat tct tct gga	816		
Asn Ser Ala Glu Val Ile Gln Ser Thr Leu Leu Asn Asn Ser Ser Gly			
260 265 270			
gag tat att gac cct cgg acg aat gga atg ttc ttg cca aat ggc cag	864		
Glu Tyr Ile Asp Pro Arg Thr Asn Gly Met Phe Leu Pro Asn Gly Gln			
275 280 285			
cta tac aac agg gac tcg agt ttt cag tcc cat ttg aat tca ttt gag	912		
Leu Tyr Asn Arg Asp Ser Ser Phe Gln Ser His Leu Asn Ser Phe Glu			
290 295 300			
gct acc tct ggt atg gca cct ctt cta gat aat gag aag gag gag tac	960		
Ala Thr Ser Gly Met Ala Pro Leu Leu Asp Asn Glu Lys Glu Glu Tyr			
305 310 315 320			
att gaa atg aat gat ctt ctg atc cct gag ctg ggt gct tct tca aca	1008		
Ile Glu Met Asn Asp Leu Leu Ile Pro Glu Leu Gly Ala Ser Ser Thr			
325 330 335			
gag aaa tcc aca gag ttc ttg aac cat ggt gaa ttt ggt gat gtt aat	1056		
Glu Lys Ser Thr Glu Phe Leu Asn His Gly Glu Phe Gly Asp Val Asn			
340 345 350			
gaa tac gac caa ttg ttc aat gac ata tct gtt ttt cag gga act tct	1104		
Glu Tyr Asp Gln Leu Phe Asn Asp Ile Ser Val Phe Gln Gly Thr Ser			
355 360 365			
aca gat ctg tct tgt ctg agt aat ttt act aat aac aca tca ggc caa	1152		
Thr Asp Leu Ser Cys Leu Ser Asn Phe Thr Asn Asn Thr Ser Gly Gln			
370 375 380			
aga cag caa tta ctt tat gaa cag ttc cag tac cag aca cct gag aac	1200		
Arg Gln Gln Leu Leu Tyr Glu Gln Phe Gln Tyr Gln Thr Pro Glu Asn			
385 390 395 400			
cag ctt aat aac tac atg cat cct agt acc act ctt aat cag ttc act	1248		
Gln Leu Asn Asn Tyr Met His Pro Ser Thr Thr Leu Asn Gln Phe Thr			
405 410 415			
gac aat atg tgg ttt aaa gat gat cag gct gct ctg tat gtt caa cca	1296		
Asp Asn Met Trp Phe Lys Asp Asp Gln Ala Ala Leu Tyr Val Gln Pro			
420 425 430			
cca caa tct tct tct gga gca ttc act tca cag tca aca ggt gtg atg	1344		
Pro Gln Ser Ser Ser Gly Ala Phe Thr Ser Gln Ser Thr Gly Val Met			
435 440 445			
cct gag tct atg aat ccc act atg agt gta aat ccc caa tac aag gaa	1392		
Pro Glu Ser Met Asn Pro Thr Met Ser Val Asn Pro Gln Tyr Lys Glu			
450 455 460			
gga caa aat ggt ggt gga aca agg agc cag ttc tca tca gct ctg tgg	1440		
Gly Gln Asn Gly Gly Thr Arg Ser Gln Phe Ser Ser Ala Leu Trp			
465 470 475 480			

```

gaa tta ttg gaa tca ata cca tca aca cca gcc tct gcc tgt gag ggt 1488
Glu Leu Leu Glu Ser Ile Pro Ser Thr Pro Ala Ser Ala Cys Glu Gly
      485                      490                      495

cct ctt aac cag acc ttt gtg cgt atg tct agc ttc agc cgc atc agg 1536
Pro Leu Asn Gln Thr Phe Val Arg Met Ser Ser Phe Ser Arg Ile Arg
      500                      505                      510

ttc aat gga acg tca gtg act agt aga aaa gtc act gta gca aag aag 1584
Phe Asn Gly Thr Ser Val Thr Ser Arg Lys Val Thr Val Ala Lys Lys
      515                      520                      525

cgt atc agt aac aga ggt ttt ctt ctg tta tca att atg ggt gct ttg 1632
Arg Ile Ser Asn Arg Gly Phe Leu Leu Leu Ser Ile Met Gly Ala Leu
      530                      535                      540

tgt gct atc ttc tgg gtg ttc aaa gcc acc gtt gga gtt atg gga aga 1680
Cys Ala Ile Phe Trp Val Phe Lys Ala Thr Val Gly Val Met Gly Arg
      545                      550                      555                      560

cct ctc ttg tcg tga 1695
Pro Leu Leu Ser *
```

&lt;210&gt; 240

&lt;211&gt; 564

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (17)...(188)

&lt;223&gt; Conserved domain

&lt;400&gt; 240

```

Met Val Asp Ser Ser Arg Asp Ser Cys Phe Lys Ala Gly Lys Phe Ser
 1          5          10          15
Ala Pro Gly Phe Arg Phe His Pro Thr Asp Glu Glu Leu Val Val Tyr
      20          25          30
Tyr Leu Lys Arg Lys Ile Cys Cys Lys Lys Leu Arg Val Asn Ala Ile
      35          40          45
Gly Val Val Asp Val Tyr Lys Val Asp Pro Ser Glu Leu Pro Gly Leu
      50          55          60
Ser Met Leu Lys Thr Gly Asp Arg Gln Trp Phe Phe Phe Thr Pro Arg
      65          70          75          80
Asn Arg Lys Tyr Pro Asn Ala Ala Arg Ser Ser Arg Gly Thr Ala Thr
      85          90          95
Gly Tyr Trp Lys Ala Thr Gly Lys Asp Arg Val Ile Glu Tyr Asn Ser
      100         105         110
Arg Ser Val Gly Leu Lys Lys Thr Leu Val Phe Tyr Arg Gly Arg Ala
      115         120         125
Pro Asn Gly Glu Arg Thr Asp Trp Val Met His Glu Tyr Thr Met Asp
      130         135         140
Glu Glu Glu Leu Gly Arg Cys Lys Asn Ala Lys Glu Tyr Tyr Ala Leu
      145         150         155         160
```

Tyr Lys Leu Tyr Lys Lys Ser Gly Ala Gly Pro Lys Asn Gly Glu G  
 165 170 175  
 Tyr Gly Ala Pro Phe Gln Glu Glu Glu Trp Val Asp Ser Asp Ser Glu  
 180 185 190  
 Asp Ala Asp Ser Val Ala Val Pro Asp Tyr Pro Val Val Arg Tyr Glu  
 195 200 205  
 Asn Gly Pro Cys Val Asp Asp Thr Lys Phe Cys Asn Pro Val Lys Leu  
 210 215 220  
 Gln Leu Glu Asp Ile Glu Lys Leu Leu Asn Glu Ile Pro Asp Ala Pro  
 225 230 235 240  
 Gly Val Asn Gln Arg Gln Phe Asp Glu Phe Val Gly Val Pro Gln Gly  
 245 250 255  
 Asn Ser Ala Glu Val Ile Gln Ser Thr Leu Leu Asn Asn Ser Ser Gly  
 260 265 270  
 Glu Tyr Ile Asp Pro Arg Thr Asn Gly Met Phe Leu Pro Asn Gly Gln  
 275 280 285  
 Leu Tyr Asn Arg Asp Ser Ser Phe Gln Ser His Leu Asn Ser Phe Glu  
 290 295 300  
 Ala Thr Ser Gly Met Ala Pro Leu Leu Asp Asn Glu Lys Glu Glu Tyr  
 305 310 315 320  
 Ile Glu Met Asn Asp Leu Leu Ile Pro Glu Leu Gly Ala Ser Ser Thr  
 325 330 335  
 Glu Lys Ser Thr Glu Phe Leu Asn His Gly Glu Phe Gly Asp Val Asn  
 340 345 350  
 Glu Tyr Asp Gln Leu Phe Asn Asp Ile Ser Val Phe Gln Gly Thr Ser  
 355 360 365  
 Thr Asp Leu Ser Cys Leu Ser Asn Phe Thr Asn Asn Thr Ser Gly Gln  
 370 375 380  
 Arg Gln Gln Leu Leu Tyr Glu Gln Phe Gln Tyr Gln Thr Pro Glu Asn  
 385 390 395 400  
 Gln Leu Asn Asn Tyr Met His Pro Ser Thr Thr Leu Asn Gln Phe Thr  
 405 410 415  
 Asp Asn Met Trp Phe Lys Asp Asp Gln Ala Ala Leu Tyr Val Gln Pro  
 420 425 430  
 Pro Gln Ser Ser Ser Gly Ala Phe Thr Ser Gln Ser Thr Gly Val Met  
 435 440 445  
 Pro Glu Ser Met Asn Pro Thr Met Ser Val Asn Pro Gln Tyr Lys Glu  
 450 455 460  
 Gly Gln Asn Gly Gly Gly Thr Arg Ser Gln Phe Ser Ser Ala Leu Trp  
 465 470 475 480  
 Glu Leu Leu Glu Ser Ile Pro Ser Thr Pro Ala Ser Ala Cys Glu Gly  
 485 490 495  
 Pro Leu Asn Gln Thr Phe Val Arg Met Ser Ser Phe Ser Arg Ile Arg  
 500 505 510  
 Phe Asn Gly Thr Ser Val Thr Ser Arg Lys Val Thr Val Ala Lys Lys  
 515 520 525  
 Arg Ile Ser Asn Arg Gly Phe Leu Leu Leu Ser Ile Met Gly Ala Leu  
 530 535 540  
 Cys Ala Ile Phe Trp Val Phe Lys Ala Thr Val Gly Val Met Gly Arg  
 545 550 555 560  
 Pro Leu Leu Ser

&lt;210&gt; 241

&lt;211&gt; 2104

&lt;212&gt; DNA



&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (93)...(1874)

&lt;400&gt; 241

```

tcgatcttta accaaatcca gttgataagg tctcttcggt gattagcaga gatctcttta 60
atttgtgaat ttcaattcat cggaacctgt tg atg gac acc acc att gat gga 113
                               Met Asp Thr Thr Ile Asp Gly
                               1           5

ttc gcc gat tct tat gaa atc agc agc act agt ttc gtc gct acc gat 161
Phe Ala Asp Ser Tyr Glu Ile Ser Ser Thr Ser Phe Val Ala Thr Asp
      10           15           20

aac acc gac tcc tct att gtt tat ctg gcc gcc gaa caa gta ctc acc 209
Asn Thr Asp Ser Ser Ile Val Tyr Leu Ala Ala Glu Gln Val Leu Thr
      25           30           35

gga cct gat gta tct gct ctg caa ttg ctc tcc aac agc ttc gaa tcc 257
Gly Pro Asp Val Ser Ala Leu Gln Leu Leu Ser Asn Ser Phe Glu Ser
      40           45           50           55

gtc ttt gac tcg ccg gat gat ttc tac agc gac gct aag ctt gtt ctc 305
Val Phe Asp Ser Pro Asp Asp Phe Tyr Ser Ser Ala Lys Leu Val Leu
      60           65           70

tcc gac ggc cgg gaa gtt tct ttc cac cgg tgc gtt ttg tca gcg aga 353
Ser Asp Gly Arg Glu Val Ser Phe His Arg Cys Val Leu Ser Ala Arg
      75           80           85

agc tct ttc ttc aag agc gct tta gcc gcc gct aag aag gag aaa gac 401
Ser Ser Phe Phe Lys Ser Ala Leu Ala Ala Lys Lys Glu Lys Asp
      90           95          100

tcc aac aac acc gcc gcc gtg aag ctc gag ctt aag gag att gcc aag 449
Ser Asn Asn Thr Ala Ala Val Lys Leu Glu Leu Lys Glu Ile Ala Lys
     105           110          115

gat tac gaa gtc ggt ttc gat tcg gtt gtg act gtt ttg gct tat gtt 497
Asp Tyr Glu Val Gly Phe Asp Ser Val Val Thr Val Leu Ala Tyr Val
    120           125          130          135

tac agc agc aga gtg aga ccg ccg cct aaa gga gtt tct gaa tgc gca 545
Tyr Ser Ser Arg Val Arg Pro Pro Pro Lys Gly Val Ser Glu Cys Ala
     140           145          150

gac gag aat tgc tgc cac gtg gct tgc cgg ccg gcg gtg gat ttc atg 593
Asp Glu Asn Cys Cys His Val Ala Cys Arg Pro Ala Val Asp Phe Met
     155           160          165

ttg gag gtt ctc tat ttg gct ttc atc ttc aag atc cct gaa tta att 641
Leu Glu Val Leu Tyr Leu Ala Phe Ile Phe Lys Ile Pro Glu Leu Ile
     170           175          180

act ctc tat cag agg cac tta ttg gac gtt gta gac aaa gtt gtt ata 689

```

Thr	Leu	Tyr	Gln	Arg	His	Leu	Leu	Asp	Val	Val	Asp	Lys	Val	Val	I	
185						190					195					
gag	gac	aca	ttg	gtt	ata	ctc	aag	ctt	gct	aat	ata	tgt	ggt	aaa	gct	737
Glu	Asp	Thr	Leu	Val	Ile	Leu	Lys	Leu	Ala	Asn	Ile	Cys	Gly	Lys	Ala	
200					205				210					215		
tgt	atg	aag	cta	ttg	gat	aga	tgt	aaa	gag	att	att	gtc	aag	tct	aat	785
Cys	Met	Lys	Leu	Leu	Asp	Arg	Cys	Lys	Glu	Ile	Ile	Val	Lys	Ser	Asn	
				220					225					230		
gta	gat	atg	gtt	agt	ctt	gaa	aag	tca	ttg	ccg	gaa	gag	ctt	gtt	aaa	833
Val	Asp	Met	Val	Ser	Leu	Glu	Lys	Leu	Pro	Glu	Glu	Leu	Val	Lys		
			235					240					245			
gag	ata	att	gat	aga	cgt	aaa	gag	ctt	ggg	ttg	gag	gta	cct	aaa	gta	881
Glu	Ile	Ile	Asp	Arg	Arg	Lys	Glu	Leu	Gly	Leu	Glu	Val	Pro	Lys	Val	
	250					255						260				
aag	aaa	cat	gtc	tcg	aat	gta	cat	aag	gca	ctt	gac	tcg	gat	gat	att	929
Lys	Lys	His	Val	Ser	Asn	Val	His	Lys	Ala	Leu	Asp	Ser	Asp	Asp	Ile	
	265					270					275					
gag	tta	gtc	aag	ttg	ctt	ttg	aaa	gag	gat	cac	acc	aat	cta	gat	gat	977
Glu	Leu	Val	Lys	Leu	Leu	Leu	Lys	Glu	Asp	His	Thr	Asn	Leu	Asp	Asp	
280					285				290					295		
gcg	tgt	gct	ctt	cat	ttc	gct	gtt	gca	tat	tgc	aat	gtg	aag	acc	gca	1025
Ala	Cys	Ala	Leu	His	Phe	Ala	Val	Ala	Tyr	Cys	Asn	Val	Lys	Thr	Ala	
				300					305					310		
aca	gat	ctt	tta	aaa	ctt	gat	ctt	gcc	gat	gtc	aac	cat	agg	aat	ccg	1073
Thr	Asp	Leu	Leu	Lys	Leu	Asp	Leu	Ala	Asp	Val	Asn	His	Arg	Asn	Pro	
			315					320					325			
agg	gga	tat	acg	gtg	ctt	cat	gtt	gct	gcg	atg	cgg	aag	gag	cca	caa	1121
Arg	Gly	Tyr	Thr	Val	Leu	His	Val	Ala	Ala	Met	Arg	Lys	Glu	Pro	Gln	
		330					335					340				
ttg	ata	cta	tct	cta	ttg	gaa	aaa	ggg	gca	agt	gca	tca	gaa	gca	act	1169
Leu	Ile	Leu	Ser	Leu	Leu	Glu	Lys	Gly	Ala	Ser	Ala	Ser	Glu	Ala	Thr	
	345					350					355					
ttg	gaa	ggg	aga	acc	gca	ctc	atg	atc	gca	aaa	caa	gcc	act	atg	gcg	1217
Leu	Glu	Gly	Arg	Thr	Ala	Leu	Met	Ile	Ala	Lys	Gln	Ala	Thr	Met	Ala	
360					365				370					375		
gtt	gaa	tgt	aat	aat	atc	ccg	gag	caa	tgc	aag	cat	tct	ctc	aaa	ggc	1265
Val	Glu	Cys	Asn	Asn	Ile	Pro	Glu	Gln	Cys	Lys	His	Ser	Leu	Lys	Gly	
				380					385					390		
cga	cta	tgt	gta	gaa	ata	cta	gag	caa	gaa	gac	aaa	cga	gaa	caa	att	1313
Arg	Leu	Cys	Val	Glu	Ile	Leu	Glu	Gln	Glu	Asp	Lys	Arg	Glu	Gln	Ile	
			395					400					405			
cct	aga	gat	gtt	cct	ccc	tct	ttt	gca	gtg	gcg	gcc	gat	gaa	ttg	aag	1361
Pro	Arg	Asp	Val	Pro	Pro	Ser	Phe	Ala	Val	Ala	Ala	Asp	Glu	Leu	Lys	

410	415	420	
atg acg ctg ctc gat ctt gaa aat aga gtt gca ctt gct caa cgt ctt			1409
Met Thr Leu Leu Asp Leu Glu Asn Arg Val Ala Leu Ala Gln Arg Leu			
425	430	435	
ttt cca acg gaa gca caa gct gca atg gag atc gcc gaa atg aag gga			1457
Phe Pro Thr Glu Ala Gln Ala Ala Met Glu Ile Ala Glu Met Lys Gly			
440	445	450	455
aca tgt gag ttc ata gtg act agc ctc gag cct gac cgt ctc act ggt			1505
Thr Cys Glu Phe Ile Val Thr Ser Leu Glu Pro Asp Arg Leu Thr Gly			
	460	465	470
acg aag aga aca tca ccg ggt gta aag ata gca cct ttc aga atc cta			1553
Thr Lys Arg Thr Ser Pro Gly Val Lys Ile Ala Pro Phe Arg Ile Leu			
	475	480	485
gaa gag cat caa agt aga cta aaa gcg ctt tct aaa acc gtg gaa ctc			1601
Glu Glu His Gln Ser Arg Leu Lys Ala Leu Ser Lys Thr Val Glu Leu			
	490	495	500
ggg aaa cga ttc ttc ccg cgc tgt tcg gca gtg ctc gac cag att atg			1649
Gly Lys Arg Phe Phe Pro Arg Cys Ser Ala Val Leu Asp Gln Ile Met			
	505	510	515
aac tgt gag gac ttg act caa ctg gct tgc gga gaa gac gac act gct			1697
Asn Cys Glu Asp Leu Thr Gln Leu Ala Cys Gly Glu Asp Asp Thr Ala			
	520	525	530
gag aaa cga cta caa aag aag caa agg tac atg gaa ata caa gag aca			1745
Glu Lys Arg Leu Gln Lys Lys Gln Arg Tyr Met Glu Ile Gln Glu Thr			
	540	545	550
cta aag aag gcc ttt agt gag gac aat ttg gaa tta gga aat tcg tcc			1793
Leu Lys Lys Ala Phe Ser Glu Asp Asn Leu Glu Leu Gly Asn Ser Ser			
	555	560	565
ctg aca gat tcg act tct tcc aca tcg aaa tca acc ggt gga aag agg			1841
Leu Thr Asp Ser Thr Ser Ser Thr Ser Lys Ser Thr Gly Gly Lys Arg			
	570	575	580
tct aac cgt aaa ctc tct cat cgt cgt cgg tga gactcttgcc tcttagtgta			1894
Ser Asn Arg Lys Leu Ser His Arg Arg Arg *			
	585	590	
atttttgctg taccatataa ttctgttttc atgatgactg taactgttta tgtctatcgt			1954
tggcgtcata tagtttcgct cttcgttttg catcctgtgt attattgctg caggtgtgct			2014
tcaaacaat gttgtaacaa tttgaaccaa tgggtatacag atttgtaata tatatttatg			2074
tacatcaaca ataaaaaaaa aaaaaaaaaa			2104
<210> 242			
<211> 593			
<212> PRT			
<213> Arabidopsis thaliana			
<220>			

&lt;221&gt; DOMAIN

&lt;222&gt; (2)...(593)

&lt;223&gt; Conserved domain

&lt;400&gt; 242

```

Met Asp Thr Thr Ile Asp Gly Phe Ala Asp Ser Tyr Glu Ile Ser Ser
 1           5           10          15
Thr Ser Phe Val Ala Thr Asp Asn Thr Asp Ser Ser Ile Val Tyr Leu
      20           25           30
Ala Ala Glu Gln Val Leu Thr Gly Pro Asp Val Ser Ala Leu Gln Leu
      35           40           45
Leu Ser Asn Ser Phe Glu Ser Val Phe Asp Ser Pro Asp Asp Phe Tyr
      50           55           60
Ser Asp Ala Lys Leu Val Leu Ser Asp Gly Arg Glu Val Ser Phe His
      65           70           75           80
Arg Cys Val Leu Ser Ala Arg Ser Ser Phe Phe Lys Ser Ala Leu Ala
      85           90           95
Ala Ala Lys Lys Glu Lys Asp Ser Asn Asn Thr Ala Ala Val Lys Leu
      100          105          110
Glu Leu Lys Glu Ile Ala Lys Asp Tyr Glu Val Gly Phe Asp Ser Val
      115          120          125
Val Thr Val Leu Ala Tyr Val Tyr Ser Ser Arg Val Arg Pro Pro Pro
      130          135          140
Lys Gly Val Ser Glu Cys Ala Asp Glu Asn Cys Cys His Val Ala Cys
      145          150          155          160
Arg Pro Ala Val Asp Phe Met Leu Glu Val Leu Tyr Leu Ala Phe Ile
      165          170          175
Phe Lys Ile Pro Glu Leu Ile Thr Leu Tyr Gln Arg His Leu Leu Asp
      180          185          190
Val Val Asp Lys Val Val Ile Glu Asp Thr Leu Val Ile Leu Lys Leu
      195          200          205
Ala Asn Ile Cys Gly Lys Ala Cys Met Lys Leu Leu Asp Arg Cys Lys
      210          215          220
Glu Ile Ile Val Lys Ser Asn Val Asp Met Val Ser Leu Glu Lys Ser
      225          230          235          240
Leu Pro Glu Glu Leu Val Lys Glu Ile Ile Asp Arg Arg Lys Glu Leu
      245          250          255
Gly Leu Glu Val Pro Lys Val Lys Lys His Val Ser Asn Val His Lys
      260          265          270
Ala Leu Asp Ser Asp Asp Ile Glu Leu Val Lys Leu Leu Leu Lys Glu
      275          280          285
Asp His Thr Asn Leu Asp Asp Ala Cys Ala Leu His Phe Ala Val Ala
      290          295          300
Tyr Cys Asn Val Lys Thr Ala Thr Asp Leu Leu Lys Leu Asp Leu Ala
      305          310          315          320
Asp Val Asn His Arg Asn Pro Arg Gly Tyr Thr Val Leu His Val Ala
      325          330          335
Ala Met Arg Lys Glu Pro Gln Leu Ile Leu Ser Leu Leu Glu Lys Gly
      340          345          350
Ala Ser Ala Ser Glu Ala Thr Leu Glu Gly Arg Thr Ala Leu Met Ile
      355          360          365
Ala Lys Gln Ala Thr Met Ala Val Glu Cys Asn Asn Ile Pro Glu Gln
      370          375          380
Cys Lys His Ser Leu Lys Gly Arg Leu Cys Val Glu Ile Leu Glu Gln
      385          390          395          400
Glu Asp Lys Arg Glu Gln Ile Pro Arg Asp Val Pro Pro Ser Phe Ala
      405          410          415

```



Ser	Arg	Thr	Ser	Ser	Phe	His	Met	Glu	Pro	Val	Lys	Asn	Asn	Gly	His	
70						75					80					
agc	aga	gcc	att	acc	ttg	cag	aac	aag	aga	aaa	cca	gag	ggg	aag	aca	345
Ser	Arg	Ala	Ile	Thr	Leu	Gln	Asn	Lys	Arg	Lys	Pro	Glu	Gly	Lys	Thr	
85					90					95					100	
gaa	aag	aga	gag	aag	aag	aag	atc	aaa	gca	gag	gat	gaa	aca	gag	cca	393
Glu	Lys	Arg	Glu	Lys	Lys	Lys	Ile	Lys	Ala	Glu	Asp	Glu	Thr	Glu	Pro	
				105					110					115		
agc	atg	aaa	ggg	aaa	tca	aac	atg	agt	aac	aca	gag	aca	tct	tca	gaa	441
Ser	Met	Lys	Gly	Lys	Ser	Asn	Met	Ser	Asn	Thr	Glu	Thr	Ser	Ser	Glu	
			120					125					130			
att	cag	aaa	cca	gat	tac	att	cat	gtt	agg	gct	aga	cga	ggg	gaa	gcc	489
Ile	Gln	Lys	Pro	Asp	Tyr	Ile	His	Val	Arg	Ala	Arg	Arg	Gly	Glu	Ala	
		135					140					145				
acc	gac	aga	cat	agc	tta	gca	gag	agg	gca	aga	aga	gaa	aag	ata	agc	537
Thr	Asp	Arg	His	Ser	Leu	Ala	Glu	Arg	Ala	Arg	Arg	Glu	Lys	Ile	Ser	
		150				155					160					
aag	aag	atg	aaa	tgt	cta	caa	gat	att	gtt	cct	gga	tgc	aac	aaa	gtt	585
Lys	Lys	Met	Lys	Cys	Leu	Gln	Asp	Ile	Val	Pro	Gly	Cys	Asn	Lys	Val	
165					170					175					180	
act	gga	aaa	gct	ggg	atg	ctt	gat	gag	atc	atc	aac	tat	gtc	caa	tct	633
Thr	Gly	Lys	Ala	Gly	Met	Leu	Asp	Glu	Ile	Ile	Asn	Tyr	Val	Gln	Ser	
				185					190					195		
ctg	caa	caa	caa	gtc	gag	ttc	ttg	tcg	atg	aaa	ctc	tct	gtc	ata	aat	681
Leu	Gln	Gln	Gln	Val	Glu	Phe	Leu	Ser	Met	Lys	Leu	Ser	Val	Ile	Asn	
			200					205					210			
cca	gaa	ctt	gag	tgt	cat	atc	gat	gat	tta	tcc	gca	aaa	cag	ttt	cag	729
Pro	Glu	Leu	Glu	Cys	His	Ile	Asp	Asp	Leu	Ser	Ala	Lys	Gln	Phe	Gln	
		215					220					225				
gct	tac	ttc	aca	ggg	cct	cca	gaa	ggg	gac	tcg	aag	cag	tca	atc	atg	777
Ala	Tyr	Phe	Thr	Gly	Pro	Pro	Glu	Gly	Asp	Ser	Lys	Gln	Ser	Ile	Met	
		230				235					240					
gcg	gat	ttt	cgg	tct	ttt	cca	tta	cat	cag	caa	gga	tct	tta	gat	tac	825
Ala	Asp	Phe	Arg	Ser	Phe	Pro	Leu	His	Gln	Gln	Gly	Ser	Leu	Asp	Tyr	
245					250					255					260	
tca	gtc	ata	aac	tca	gac	cac	acc	aca	tct	ctc	ggc	gct	aaa	gat	cat	873
Ser	Val	Ile	Asn	Ser	Asp	His	Thr	Thr	Ser	Leu	Gly	Ala	Lys	Asp	His	
			265						270				275			
aca	tca	tca	agc	tgg	gaa	act	cac	tca	cag	tgt	ctt	tac	aac	agc	ttg	921
Thr	Ser	Ser	Ser	Trp	Glu	Thr	His	Ser	Gln	Cys	Leu	Tyr	Asn	Ser	Leu	
			280					285					290			
aga	acc	gat	tct	gtt	tcc	aat	ttc	ttc	agc	ctc	aag	taa	aaaaa	attagg		970
Arg	Thr	Asp	Ser	Val	Ser	Asn	Phe	Phe	Ser	Leu	Lys	*				

295

300

gatagcctca ttaaaaaaat cgcggttttt tgttggtgtc ttatccattt atctatctta 1030  
 tctgaaattt gaaccagaaa gacagaggaa accaatccaa agatctttct caatctatta 1090  
 tcttcataca aatatagtga ttacatata ttccagggga tatgtatatg tgtagaagaa 1150  
 agagaaaaaa ctcttggtgt catagcaatt cctttttttg tacattgtag aatcaaactc 1210  
 ttgtggtcgt aacaattatt tccttcacaa attacaacta cacttgattt aatggagatg 1270  
 ccttttggcc ctgggatcaa ca 1292

&lt;210&gt; 244

&lt;211&gt; 304

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; DOMAIN

&lt;222&gt; (140)...(204)

&lt;223&gt; Conserved domain

&lt;400&gt; 244

Met Asp Leu Ser Val Leu Asp Arg Leu Lys Trp Leu Gln Gln Gln Gln  
 1 5 10 15  
 Met Val Ser Pro Glu Phe Leu Gln Ile Leu Gly Ser Asp Gly Arg Glu  
 20 25 30  
 Glu Leu Lys Arg Val Glu Ser Tyr Leu Gly Asn Asn Asn Asp Glu Leu  
 35 40 45  
 Gln Ser Phe Arg His Phe Pro Glu Phe Gly Pro Asp Tyr Asp Thr Thr  
 50 55 60  
 Asp Gly Cys Ile Ser Arg Thr Ser Ser Phe His Met Glu Pro Val Lys  
 65 70 75 80  
 Asn Asn Gly His Ser Arg Ala Ile Thr Leu Gln Asn Lys Arg Lys Pro  
 85 90 95  
 Glu Gly Lys Thr Glu Lys Arg Glu Lys Lys Lys Ile Lys Ala Glu Asp  
 100 105 110  
 Glu Thr Glu Pro Ser Met Lys Gly Lys Ser Asn Met Ser Asn Thr Glu  
 115 120 125  
 Thr Ser Ser Glu Ile Gln Lys Pro Asp Tyr Ile His Val Arg Ala Arg  
 130 135 140  
 Arg Gly Glu Ala Thr Asp Arg His Ser Leu Ala Glu Arg Ala Arg Arg  
 145 150 155 160  
 Glu Lys Ile Ser Lys Lys Met Lys Cys Leu Gln Asp Ile Val Pro Gly  
 165 170 175  
 Cys Asn Lys Val Thr Gly Lys Ala Gly Met Leu Asp Glu Ile Ile Asn  
 180 185 190  
 Tyr Val Gln Ser Leu Gln Gln Gln Val Glu Phe Leu Ser Met Lys Leu  
 195 200 205  
 Ser Val Ile Asn Pro Glu Leu Glu Cys His Ile Asp Asp Leu Ser Ala  
 210 215 220  
 Lys Gln Phe Gln Ala Tyr Phe Thr Gly Pro Pro Glu Gly Asp Ser Lys  
 225 230 235 240  
 Gln Ser Ile Met Ala Asp Phe Arg Ser Phe Pro Leu His Gln Gln Gly  
 245 250 255  
 Ser Leu Asp Tyr Ser Val Ile Asn Ser Asp His Thr Thr Ser Leu Gly  
 260 265 270  
 Ala Lys Asp His Thr Ser Ser Ser Trp Glu Thr His Ser Gln Cys Leu  
 275 280 285  
 Tyr Asn Ser Leu Arg Thr Asp Ser Val Ser Asn Phe Phe Ser Leu Lys

290

295

300

&lt;210&gt; 245

&lt;211&gt; 1231

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (107)...(1084)

&lt;400&gt; 245

tatagagaga ggaaaaaact atggctcaca taaagatcta acagagacag tagtatagtc 60  
 ccacttcccc ccaccaaag cttcgatcat catcatcatc atcatc atg gat ecg 115  
 Met Asp Pro  
 1

gat ccg gat cat aac cat cga ccc aac ttc cct ctc cag ctt ctt gat 163  
 Asp Pro Asp His Asn His Arg Pro Asn Phe Pro Leu Gln Leu Leu Asp  
 5 10 15

tct tct acc tcc tcc tct tcc act tcc tta gcc atc atc tct act act 211  
 Ser Ser Thr Ser Ser Ser Ser Thr Ser Leu Ala Ile Ile Ser Thr Thr  
 20 25 30 35

tcc gaa cct aac tcc gaa cct aag aag cct cct cct aaa cga acc tct 259  
 Ser Glu Pro Asn Ser Glu Pro Lys Lys Pro Pro Pro Lys Arg Thr Ser  
 40 45 50

act aaa gac cga cac acc aaa gtc gaa ggc cga ggc cgt cgg atc cgt 307  
 Thr Lys Asp Arg His Thr Lys Val Glu Gly Arg Gly Arg Arg Ile Arg  
 55 60 65

atg cct gcc atg tgt gct gca cgt gtc ttt cag ctc aca cgt gag ctt 355  
 Met Pro Ala Met Cys Ala Ala Arg Val Phe Gln Leu Thr Arg Glu Leu  
 70 75 80

ggc cac aaa tcc gac ggt gaa act att gag tgg cta ctc caa caa gca 403  
 Gly His Lys Ser Asp Gly Glu Thr Ile Glu Trp Leu Leu Gln Gln Ala  
 85 90 95

gaa cca gcg gtt ata gcc gct aca ggg act gga acc att ccg gct aac 451  
 Glu Pro Ala Val Ile Ala Ala Thr Gly Thr Gly Thr Ile Pro Ala Asn  
 100 105 110 115

ttc act tct tta aac atc tca ctt cgt agc tca aga tct tct ctc tct 499  
 Phe Thr Ser Leu Asn Ile Ser Leu Arg Ser Ser Arg Ser Ser Leu Ser  
 120 125 130

gct gct cat ctt cgt aca act cct agt agc tat tac ttt cat tca cca 547  
 Ala Ala His Leu Arg Thr Thr Pro Ser Ser Tyr Tyr Phe His Ser Pro  
 135 140 145

cat cag tcc atg act cat cat ctt caa cat cag cat cag gtt cgt ccc 595  
 His Gln Ser Met Thr His His Leu Gln His Gln His Gln Val Arg Pro  
 150 155 160



```

aag aac gag tca cat tct tcg tct tct tct tct tca cag ctt tta gat 643
Lys Asn Glu Ser His Ser Ser Ser Ser Ser Ser Ser Ser Gln Leu Leu Asp
165 170 175

cac aac caa atg ggt aac tat cta gta caa tca act gct gga tct tta 691
His Asn Gln Met Gly Asn Tyr Leu Val Gln Ser Thr Ala Gly Ser Leu
180 185 190 195

cct acg agt cag agt cct gca acg gca ccg ttt tgg agt agt ggt gac 739
Pro Thr Ser Gln Ser Pro Ala Thr Ala Pro Phe Trp Ser Ser Gly Asp
200 205 210

aac aca cag aat ctt tgg gct ttt aat att aat cct cat cat tcc ggt 787
Asn Thr Gln Asn Leu Trp Ala Phe Asn Ile Asn Pro His His Ser Gly
215 220 225

gtt gtc gcc gga gat gtt tac aac ccc aac agt ggt ggt agt ggc ggc 835
Val Val Ala Gly Asp Val Tyr Asn Pro Asn Ser Gly Gly Ser Gly Gly
230 235 240

ggt agt gga gtt cat ttg atg aat ttt gca gct cct att gct ttg ttt 883
Gly Ser Gly Val His Leu Met Asn Phe Ala Ala Pro Ile Ala Leu Phe
245 250 255

tct gga cag cct ttg gct tct ggt tat gga gga gga gga ggt ggc ggt 931
Ser Gly Gln Pro Leu Ala Ser Gly Tyr Gly Gly Gly Gly Gly Gly Gly
260 265 270 275

gga gaa cat agc cat tat gga gtt tta gcg gcg ttg aat gct gct tac 979
Gly Glu His Ser His Tyr Gly Val Leu Ala Ala Leu Asn Ala Ala Tyr
280 285 290

cga ccg gtg gcg gag acg ggg aac cat aac aac aac cag caa aac cgt 1027
Arg Pro Val Ala Glu Thr Gly Asn His Asn Asn Asn Gln Gln Asn Arg
295 300 305

gac gga gat cat cat cac aac cat caa gaa gat gga agc acc agt cat 1075
Asp Gly Asp His His His Asn His Gln Glu Asp Gly Ser Thr Ser His
310 315 320

cat tcc tag gcaaacatac acaaacaaat atattctgtg agatttattt 1124
His Ser *
325

ttcttttttt gtcctttcgt ttgtttgttt gttcttaaca agcgtgtttt ttttgcattg 1184
cttctttttt atatattttt atttatttct aaattctaata aaaaaaa 1231

<210> 246
<211> 325
<212> PRT
<213> Arabidopsis thaliana

<220>
<221> DOMAIN
<222> (51)...(114)
<223> Conserved domain

```

&lt;400&gt; 246

```

Met Asp Pro Asp Pro Asp His Asn His Arg Pro Asn Phe Pro Leu Gln
 1      5      10      15
Leu Leu Asp Ser Ser Thr Ser Ser Ser Thr Ser Leu Ala Ile Ile
 20      25      30
Ser Thr Thr Ser Glu Pro Asn Ser Glu Pro Lys Lys Pro Pro Pro Lys
 35      40      45
Arg Thr Ser Thr Lys Asp Arg His Thr Lys Val Glu Gly Arg Gly Arg
 50      55      60
Arg Ile Arg Met Pro Ala Met Cys Ala Ala Arg Val Phe Gln Leu Thr
 65      70      75      80
Arg Glu Leu Gly His Lys Ser Asp Gly Glu Thr Ile Glu Trp Leu Leu
 85      90      95
Gln Gln Ala Glu Pro Ala Val Ile Ala Ala Thr Gly Thr Gly Thr Ile
100      105      110
Pro Ala Asn Phe Thr Ser Leu Asn Ile Ser Leu Arg Ser Ser Arg Ser
115      120      125
Ser Leu Ser Ala Ala His Leu Arg Thr Thr Pro Ser Ser Tyr Tyr Phe
130      135      140
His Ser Pro His Gln Ser Met Thr His His Leu Gln His Gln His Gln
145      150      155      160
Val Arg Pro Lys Asn Glu Ser His Ser Ser Ser Ser Ser Ser Ser Gln
165      170      175
Leu Leu Asp His Asn Gln Met Gly Asn Tyr Leu Val Gln Ser Thr Ala
180      185      190
Gly Ser Leu Pro Thr Ser Gln Ser Pro Ala Thr Ala Pro Phe Trp Ser
195      200      205
Ser Gly Asp Asn Thr Gln Asn Leu Trp Ala Phe Asn Ile Asn Pro His
210      215      220
His Ser Gly Val Val Ala Gly Asp Val Tyr Asn Pro Asn Ser Gly Gly
225      230      235      240
Ser Gly Gly Gly Ser Gly Val His Leu Met Asn Phe Ala Ala Pro Ile
245      250      255
Ala Leu Phe Ser Gly Gln Pro Leu Ala Ser Gly Tyr Gly Gly Gly Gly
260      265      270
Gly Gly Gly Gly Glu His Ser His Tyr Gly Val Leu Ala Ala Leu Asn
275      280      285
Ala Ala Tyr Arg Pro Val Ala Glu Thr Gly Asn His Asn Asn Asn Gln
290      295      300
Gln Asn Arg Asp Gly Asp His His His Asn His Gln Glu Asp Gly Ser
305      310      315      320
Thr Ser His His Ser
325

```